



UNIVERSIDADE CATÓLICA PORTUGUESA

ILUSÕES SONORAS – UM ESTUDO SOBRE A APLICAÇÃO DA ILUSÃO SONORA
DA ESCALA DE SHEPARD EM COMPOSIÇÃO MUSICAL DIGITAL

Tese apresentada à Universidade Católica Portuguesa
para obtenção do grau de Doutor em *Ciência e Tecnologia das Artes, especialidade em*
Informática Musical

por

Pedro Luís de Moraes Patrício

ESCOLA DAS ARTES
(Janeiro 2015)



UNIVERSIDADE CATÓLICA PORTUGUESA

ILUSÕES SONORAS – UM ESTUDO SOBRE A APLICAÇÃO DA ILUSÃO SONORA
DA ESCALA DE SHEPARD EM COMPOSIÇÃO MUSICAL DIGITAL

Tese apresentada à Universidade Católica Portuguesa
para obtenção do grau de Doutor em *Ciência e Tecnologia das Artes, especialidade em*
Informática Musical

Por Pedro Luís de Moraes Patrício

Sob orientação de António de Sousa Dias e Luís Gustavo Martins

ESCOLA DAS ARTES

(Janeiro 2015)

RESUMO

As ilusões sonoras sempre exerceram um grande fascínio sobre os seres humanos. Destas, a que mais nos fascinou e revelou ter maior potencial musical por explorar foi a Ilusão Sonora da Escala de Shepard. Esta ilusão é uma escala musical dividida em 12 partes iguais, que ao ser escutada gera a sensação de perpetuidade sonora. Pelo facto, de ser constituída por sons ambíguos (i.e., o Som de Shepard) em termos de percepção de altura sonora (provavelmente conseguiremos identificar o nome da nota musical do som, mas teremos muita dificuldade em determinar a oitava a que pertence) permite criar novas possibilidades de geração de som e de escuta musical. Assim, esta dissertação propõe a criação de uma fonte sonora baseada no Som de Shepard com o objectivo de ser utilizada em projectos musicais. Consequentemente, criámos um método de gerar o Som de Shepard através de uma envolvente espectral alternativa, tanto quanto sabemos, nunca antes foi usada para reproduzir a respectiva ilusão, mesmo quando esta foi aplicada a uma melodia constituída por uma sequência de saltos intervalares não regulares. A melodia foi intitulada de Ilusão Sonora da Melodia Perpétua porque quando é escutada cria a sensação de perpetuidade sonora como acontece na Ilusão Sonora da Escala de Shepard. Para além disso, utilizámos a fonte sonora proposta como gerador de som na concepção de um instrumento digital multimédia e em composição musical.

PALAVRAS-CHAVE

Ilusões, ilusão sonora da escala de shepard, som de shepard, curvas de igual percepção de intensidade sonora, filtro A, composição musical digital, composição musical em tempo real, instrumento digital multimédia, interpretação musical com dispositivos electrónicos móveis, *pure data*.

ABSTRACT

The auditory illusions have always exerted a great fascination on human beings. Of these, the one that most fascinated us and revealed to have greater musical potential for exploring was the Shepard Scale Illusion. This illusion is a musical scale divided into 12 equal parts. When it is heard it creates the sensation of sonic perpetuity. Being constituted by ambiguous sounds (i.e., Shepard tone) in terms of pitch (it will probably be easy to identify the musical note of the sound, but it will be very difficult to determine the octave that belongs to), allows creating new possibilities of sound generation and musical listening. Thus, this dissertation proposes the creation of a sound source based on Shepard tone with the purpose to be used in musical projects. Consequently, we developed a method to generate the Shepard tone through an alternative spectral envelope, which as far as we know has never been used before to reproduce the Shepard Scale Illusion, even when it was applied to a melody constituted by a sequence of non-regular intervals leaps. The melody was titled Perpetual Melody Auditory Illusion because when it is heard it creates the sensation of sonic perpetuity, as is the case of the Shepard Scale Illusion. Furthermore, we used the proposal sound source as sound generator in the design of a multimedia digital instrument and musical composition.

KEYWORDS

Illusions, shepard scale illusion, shepard tone, equal-loudness curves, filter A, digital musical composition, real-time musical composition, multimedia digital instrument, mobile music performance, pure data.

A Roger N. Shepard (1929) pela magnífica ilusão sonora, a qual exerceu (quando a escutámos pela primeira vez) e continua a exercer um grande fascínio.

AGRADECIMENTOS

Gostaríamos de começar por agradecer à Fundação para a Ciência e Tecnologia, a bolsa individual de doutoramento concedida (SFRH/BD/61599/2009). O facto, de a bolsa nos ter permitido trabalhar na proposta de investigação, em regime de exclusividade possibilitou-nos alcançar resultados científicos de maneira mais profícua e eficiente.

Agradecemos também ao nosso orientador, António de Sousa Dias e co-orientador, Luís Gustavo Martins todo o apoio incondicional prestado. Enaltecemos sobretudo, o facto de nos terem proporcionado um ambiente de trabalho eficaz, que se pautou, simultaneamente, pelo respeito, pela preservação e pelo estímulo da individualidade e da livre iniciativa de investigação.

Por último, não menos importante, agradecemos a todas as pessoas, que connosco se cruzaram, quer em conferências, simpósios, seminários, concertos, etc., e cujo contributo, mesmo pontual, nos fez crescer, desenvolver, amadurecer, apurar e tornar ainda mais eficazes e sofisticadas as nossas capacidades académicas, científicas e artísticas. De entre essas pessoas, destacamos em termos gerais, os revisores anónimos das conferências com escrutínio científico, onde o nosso trabalho foi sendo publicado, e a nível particular a Eric Oña, António Ferreira, Helena Mena Matos e Sofia Lourenço, pelo facto, de terem sido pessoas preponderantes em determinados momentos do nosso percurso de investigação, no sentido, em que nos ajudaram a ultrapassar questões logísticas, científicas, bibliográficas, etc.

A TODOS, MUITO OBRIGADO.

ÍNDICE GERAL

Introdução.....	1
Principais Contribuições	3
Aplicações da Fonte Sonora	5
Estrutura da Dissertação	6
PARTE I – Termos, noções e conceitos.....	11
I Sistema Auditivo.....	13
I.1 Ouvido.....	13
I.1.1 Ouvido Externo	14
I.1.2 Ouvido Médio	17
I.1.3 Ouvido Interno	18
I.2 Cérebro.....	23
I.2.1 Nervo Auditivo	23
I.2.2 Córtex Cerebral	24
I.2.3 Células Nervosas (Neurónios)	24
I.2.4 Transmissão dos Impulsos Nervosos	25
I.2.5 Tempo de Propagação dos Impulsos Nervosos	26
II Psicoacústica.....	27
III Psicologia Cognitiva	33
IV Principais Teorias da Percepção	35
IV.1.1 Gibson e a Percepção Directa ou Processamento Ascendente.....	35
IV.1.2 Gregory e a Percepção Construtivista, Indirecta ou Processamento Descendente	38
IV.2 Principais Mecanismos da Percepção	43
IV.2.1 Inferência Inconsciente	44
IV.2.2 Percepção de Preenchimento	45

IV.2.3	Princípios Gestálticos de Agrupamento	47
V	Ilusões: Generalidades.....	53
V.1	Definição de Ilusão.....	53
V.2	Classificação das Ilusões	54
V.3	Um Breve Estudo sobre os Processos de Geração de Ilusões	56
V.4	Mecanismos de Actuação das Ilusões na Percepção Humana.....	58
PARTE II	- Ilusão Sonora da Escala de Shepard	61
VI	Ilusão Sonora da Escala de Shepard	63
VI.1	Som de Shepard.....	65
VI.1.1	Determinação Manual das Frequências Individuais	67
VI.1.2	Determinação Manual das Amplitudes Individuais	67
VI.2	Escala de Shepard	69
VI.2.1	Percepção Auditiva da Ilusão Sonora da Escala de Shepard	73
VI.2.2	Factores que Contribuem para a Realização da Ilusão Sonora da Escala de Shepard	74
VI.2.3	Som de Shepard e Ilusão Sonora da Escala de Shepard Revisitados.....	76
VII	Implementação da Ilusão Sonora da Escala de Shepard em Síntese de Som Digital	81
VII.1	Reconstituição da Ilusão Sonora da Escala de Shepard	81
VII.2	Determinação Automática das Frequências Individuais	82
VII.3	Determinação Automática das Amplitudes Individuais.....	85
VII.4	Reprodução Áudio da Ilusão Sonora da Escala de Shepard	88
VII.5	Percepção Auditiva da Reprodução Áudio da Ilusão Sonora da Escala de Shepard.....	89
VIII	Análise Espectral do Som de Shepard	91
VIII.1	Análise de Fourier	91
VIII.2	Breve Descrição de uma DFT	92
VIII.3	Transformada rápida de Fourier (FFT)	92
VIII.4	Análise do Som de Shepard através de uma transformada rápida de Fourier real (RFFT) realizada em Pure Data.....	93
PARTE III	– Desenvolvimento, aplicações e composição	99

IX	Chromasom: Uma Fonte Sonora paradoxal	101
IX.1	Determinação das Frequências Individuais do Chromasom	102
IX.2	Envolventes Espectrais Utilizadas por Outros Autores Aplicadas ao Chromasom	104
IX.2.1	Envolvente Espectral de Pollack	104
IX.2.2	Envolvente Espectral de Burns	105
IX.2.3	Envolvente Espectral de Nakajima et al.	108
IX.2.4	Envolvente Espectral de Deutsch	110
IX.3	Percepção Auditiva das Envolventes Espectrais	112
IX.4	Determinação da Envolvente Espectral Alternativa	113
IX.4.1	Curvas de Igual Percepção de Intensidade Sonora	113
IX.5	Exercícios Musicais Aplicados ao Chromasom	121
IX.5.1	Primeira Série de Exercícios Musicais	122
IX.5.2	Segunda Série de Exercícios Musicais	124
IX.5.3	Percepção Auditiva das Duas Séries de Exercícios Musicais Aplicadas ao Chromasom	126
IX.5.4	Experiência Auditivo-Musical	126
X	Aplicações do Chromasom e Composição	133
X.1	Aplicação do Chromasom em Síntese por Modulação de Frequência	133
X.1.1	Modulação de Frequência	133
X.1.2	Um Chromasom a Modular outro Chromasom	136
X.2	MuDI - Instrumento Digital Multimédia	138
X.2.1	Trabalho Relacionado	139
X.2.2	Sistema do MuDI	142
X.2.3	Manuseamento	150
X.2.4	Calibração do MuDI	151
X.2.5	Partitura Vídeo	155
X.3	Ciclo <i>Música Digital Contemporânea Erudita</i> (MuDiCE)	156
X.3.1	mudice\$1 – música de uma tecla só (a pretensiosa)	157
X.3.2	mudice\$2 – a irreversível.	163
X.3.3	mudice\$3 – a perpétua	166
X.3.4	Perpetual Melody – contrasting moments.	169
X.4	Ciclo <i>Sonificação – Filmes para Música</i>	172

Conclusões e Direcções Futuras.....	175
Bibliografia.....	181
Glossário.....	191
Anexos.....	193
Anexo 1: Publicações, Conferências.....	195
Lista de Artigos do Autor Publicados durante o Doutoramento em Conferências com Escrutínio Científico	195
Lista de Conferências com Escrutínio Científico Realizadas pelo Autor durante o Doutoramento.....	196
Anexo 2: Produção Artística	197
Obras Musicais Compostas e Estreadas Durante o Doutoramento	197
Concertos Realizados Durante o Doutoramento	197
Publicação em CD.....	198
Divulgação Radiofónica.....	198
Anexo 3: Listagem do conteúdo do DVD.....	199
Anexo 4: Formulário	203

ÍNDICE DE FIGURAS

Figura 1. Representação das estruturas do ouvido. As estruturas e os seus elementos não se encontram representados à escala real. Adap. de Mathews (1999, p. 2).	13
Figura 2. Principais estruturas do ouvido médio. Adap. de Warren (2008, p. 7).	17
Figura 3. Ilustração da cadeia dos ossos vestibulares e da ligação do estribo à janela oval. Adap. de Warren (2008, p. 8).	18
Figura 4. Corte longitudinal da cóclea. Adap. de Mathews (1999, p. 6).	19
Figura 5. Representação do Órgão de Corti. Adap. de Maroonroge, Emanuel e Letowski (2009, p. 292).	20
Figura 6. Representação cônica da cóclea. Adap. de Warren (2008, p. 8).	21
Figura 7. Distribuição das frequências sonoras ao longo da cóclea e da respectiva membrana basilar. As frequências sonoras distribuem-se de forma descendente (i.e., das altas para as baixas frequências), desde a extremidade basal até a extremidade distal da membrana basilar. Adap. de Warren (2008, p. 12).	22
Figura 8. Células nervosas (neurónios) distribuídas no córtex cerebral.	24
Figura 9. Constituição geral de um neurónio. Adap. de Mathews (1999, p. 12).	25
Figura 10. Analogia psicoacústica do tímpano.	27
Figura 11. Representação da forma de onda de um intervalo de quinta perfeita (C3-G3) tocado em uníssono. O intervalo foi obtido através de síntese de som digital de dois sons puros de 130.8 Hz e 195.9 Hz.	30
Figura 12. Representação da forma de onda de um intervalo de segunda menor (C3-Db3) tocado em uníssono. O intervalo foi obtido através de síntese de som digital de dois sons puros de 130.8 Hz e 138.5 Hz.	30
Figura 13. Combinação de duas ondas sinusoidais; f_1 é a frequência fixa e f_2 a frequência variável. Quando as duas frequências se encontram muito próximas originam batimentos (a). Dentro de uma determinada banda crítica percebemos uma sensação de rugosidade sonora (b). Dentro da área representada pela letra (c) ouvimos praticamente um único som (fusão sonora). Fora desta área, começamos a ouvir com clareza dois sons puros distintos (d). Adap. de Hall (1990, pp. 384).	31
Figura 14. Sobreposição. Adap. Bregman (1990, p. 628).	37
Figura 15. Ilusão Visual do Cubo de Necker. Adap. de Gregory (1977, p. 192).	39
Figura 16. Posição inicial do cubo (imagem da esquerda). Posição posterior do cubo (imagem da direita).	39
Figura 17. Imagem ambígua. Adap. de Goolkasian e Woodberry (2010, p. 170).	40
Figura 18. Ilusão Visual das pernas do elefante de Shepard. Adap. de Shepard (1990, p. 79).	43
Figura 19. Inferência inconsciente. Somos capazes de reconhecer uma porta, apesar de só vermos um retângulo longo e estreito quando a mesma se encontra entreaberta.	44
Figura 20. Triângulo de Kanizsa. Adap. de Kanizsa (1955, p. 16).	46
Figura 21. Princípio gestáltico de proximidade. Adap. de Palmer et al. (2003, p. 312).	48
Figura 22. Princípio gestáltico de similaridade. Adap. de Palmer et al. (2003, p. 312).	48
Figura 23. Princípio gestáltico de simetria. Adap. de Wagemans et al. (2012, p. 1180).	48
Figura 24. Princípio gestáltico de boa continuação. Adap. de Deutsch (1999, p. 300); Shepard (1999, p. 32).	49
Figura 25. Princípio gestáltico de destino comum. Adap. de Palmer (2002, p. 212).	49

Figura 26. Espiral. Adap. de Gregory (1968, p. 280).	54
Figura 27. Ilusão Visual de Ponzo. Adap. de Gregory (1968, p. 284).	56
Figura 28. Fotografia dos Carris do Caminho de Ferro. Adap. de Gregory (1968, p. 282).	56
Figura 29. Desenho de dois rectângulos brancos sobre a fotografia dos Carris do Caminho de Ferro. Adap. de Gregory (1968, p. 282).	57
Figura 30. Litografia da Ilusão Visual da Escada Interminável de Escher. Na parte superior da imagem é possível observar os monges a subir ou descer as escadas ao longo de um ciclo interminável. Adap. de Hofstadter (1979, p. 20).	63
Figura 31. Imagem produzida a partir da parte superior da litografia de Escher. Por mais que a figura humana suba ou desça as escadas permanecerá enclausurada num ciclo interminável.	64
Figura 32. O Poste de Barbeiro possui faixas vermelhas e azuis dispostas na diagonal sobre um fundo branco. Quando colocadas a girar e observadas de um determinado ângulo parecem mover-se para cima ou para baixo, ao longo de um movimento contínuo no sentido da rotação.	65
Figura 33. Som de Shepard. Na imagem estão representadas as intensidades (dB) de dez componentes separadas por intervalos de oitava simbolizadas pelas linhas verticais contínuas. As linhas verticais a tracejado correspondem a um deslocamento ascendente nas frequências de todos as componentes. As componentes das extremidades possuem sempre as amplitudes mais baixas e as componentes intermédias as amplitudes mais altas. Adap. de Shepard (1964, p. 2347).	66
Figura 34. Representação bidimensional da percepção de altura de sons complexos, cujas componentes se encontram separadas por intervalos de oitava. Na Figura, encontramos representada a dimensão vertical dos sons (altura espectral) e a dimensão circular (altura tonal). Adap. de Deutsch (2008, p. 590).	70
Figura 35. Quando as oitavas de uma melodia conhecida são distribuídas de forma aleatória (c) torna-se bastante difícil reconhecer a melodia original (a). Pelo contrário, quando se distribuem de maneira a preservar a sua envolvente melódica (b) torna-se mais fácil o reconhecimento da melodia original. Adap. de Shepard (1999, p. 161).	72
Figura 36. Representação da escala musical da Ilusão Sonora da Escala de Shepard. A imagem representa a relação circular ou horizontal dos sons complexos, ao longo da respectiva ilusão. A ilusão é uma escala musical dividida em 12 partes iguais. Adap. de Deutsch (1987, p. 563).	73
Figura 37. Implementação da Equação 2. Para $t = 1$ introduzir na entrada do lado esquerdo do objecto “ <i>expr</i> ” os valores de C (i.e., de um a dez); para $t = 2$ realizar o mesmo procedimento, e assim sucessivamente até $t = 12$.	82
Figura 38. Determinação e extracção automática das frequências individuais, sob a forma de listas.	82
Figura 39. Listas das frequências individuais obtidas a partir do procedimento representado na Figura 38.	83
Figura 40. Processo de visualização das notas musicais correspondentes às frequências individuais que se encontram representadas na Tabela 2.	84
Figura 41. Implementação das Equações 3 e 4. Para $t = 1$ introduzir na entrada do lado esquerdo do objecto “ <i>expr</i> ” os valores de C (i.e., de um a dez); para $t = 2$ realizar o mesmo procedimento, e assim sucessivamente até $t = 12$.	85
Figura 42. Determinação automática das amplitudes individuais a partir do procedimento representado na Figura 41.	85
Figura 43. Listas das amplitudes individuais obtidas a partir dos procedimentos representados na Figura 42.	86
Figura 44. Representação gráfica das envolventes de amplitude dos sons que constituem a Ilusão Sonora da Escala de Shepard.	87
Figura 45. Comparação entre as envolventes de amplitude “ <i>amp1</i> ” e “ <i>amp12</i> ”. As envolventes encontram-se sobrepostas.	87

Figura 46. Síntese do primeiro som da Ilusão Sonora da Escala de Shepard. O ponto (a) representa as frequências individuais (Hz) das componentes e o ponto (b) as respectivas amplitudes (dB). Por sua vez, os pontos (c) e (d) representam a envolvente de amplitude geral aplicada ao som. A amplitude geral do som consiste no seguinte: a amplitude sobe linearmente de zero até 0.002 em 10 milissegundos; após 120 milissegundos, desce de 0.002 até zero em 10 milissegundos.....	88
Figura 47. Interface de utilizador da reprodução áudio da Ilusão Sonora da Escala de Shepard.	89
Figura 48. Implementação áudio da Equação 5 e visualização da janela <i>Hanning</i>	94
Figura 49. Implementação áudio da Equação 6.	96
Figura 50. Espectrograma do primeiro do Som de Shepard.	97
Figura 51. Envolvente espectral de Pollack. As linhas verticais representam as dez componentes que se encontram separadas por intervalos de oitava e a linha oblíqua a envolvente espectral. Adap. de Pollack (1978, p. 202).	104
Figura 52. Implementação da Equação 7 e extracção automática das amplitudes individuais sob a forma de lista.	105
Figura 53. Representação gráfica da reprodução da envolvente espectral de Pollack.	105
Figura 54. Envolvente espectral de Burns. No eixo <i>x</i> , as frequências estão representadas de forma logarítmica. As linhas verticais simbolizam as dez componentes do som e a linha constituída por pontos a envolvente espectral. Adap. de Burns (1981, p. 468).	106
Figura 55. Conversão dos valores negativos de intensidade em positivos através do objecto “ <i>abs</i> ”. Normalização dos valores obtidos, através da Equação 8, numa escala compreendida entre zero e um.	107
Figura 56. Representação gráfica da reprodução da envolvente espectral de Burns.	107
Figura 57. Envolvente espectral de Nakajima <i>et al.</i> As linhas verticais representam as componentes do som e o tracejado a envolvente espectral. Adap. de Nakajima <i>et al.</i> , (1988, p. 7).	108
Figura 58. Conversão dos valores negativos de intensidade em positivos através do objecto “ <i>abs</i> ”. Normalização dos valores obtidos, através da Equação 8, numa escala compreendida entre zero e um.	109
Figura 59. Representação gráfica da reprodução da envolvente espectral de Nakajima <i>et al.</i>	109
Figura 60. Envolvente espectral de Deutsch. No eixo <i>x</i> , as frequências estão representadas de forma logarítmica. As linhas verticais representam as seis componentes separadas por intervalos de oitava que constituem o som e a linha curva a envolvente espectral. Adap. de Deutsch et al. (1986, p. 1348).	110
Figura 61. Implementação da Equação 9 e extracção automática das listas das amplitudes individuais.	111
Figura 62. Representação gráfica da reprodução da envolvente espectral de Deutsch.	112
Figura 63. Interface de utilizador “ <i>envolventes_espectrais</i> ” usada para escutar o resultado sonoro da reprodução da Ilusão Sonora da Escala de Shepard realizada através de envolventes espectrais de outros autores.	112
Figura 64. Comparação entre as novas curvas propostas por Suzuki e Takeshima e as curvas de Robinson e Dadson. Adap. de Suzuki e Takeshima (2004, p. 928).	115
Figura 65. Comparação entre as novas curvas propostas por Suzuki e Takeshima e as curvas de Fletcher e Munson. Adap. de Suzuki e Takeshima (2004, p. 928).	116
Figura 66. Filtro A. Nos 1000 Hz (frequência de referência) o nível de pressão sonora é igual a zero dB. Adap. de Wong (2007, p. 458).	117
Figura 67. Implementação e extracção automática das listas das amplitudes individuais obtidas a partir da Equação 10.	118
Figura 68. Representação gráfica da envolvente espectral alternativa. As linhas verticais representam nove componentes separadas por intervalos de oitava.	118
Figura 69. Comparação entre as curvas invertidas de 40 <i>phons</i> de Fletcher e Munson e de Suzuki e Takeshima (a) com a curva do Filtro A (b) e a envolvente espectral alternativa	

obtida para o Chromasom (c). Para que todas as curvas apresentassem a mesma posição, de maneira a facilitar a comparação, invertamos a imagem do canto esquerdo....	119
Figura 70. Listas dos valores obtidos através do procedimento representado na Figura 67.	120
Figura 71. Interface de utilizador “Chromasom”.	121
Figura 72. Implementação da Equação 11 (a) e Equação 12 (b).	123
Figura 73. Determinação automática das frequências individuais das escalas musicais divididas em 24 e 48 partes iguais.	123
Figura 74. Determinação automática das amplitudes individuais das escalas musicais divididas em 24 e 48 partes iguais, obtidas a partir da Equação 10.	124
Figura 75. Interface de utilizador “series_exercicios”. Esta interface permite escutar os resultados sonoros obtidos a partir das duas séries de exercícios musicais aplicados ao Chromasom.	125
Figura 76. <i>The Endless Stairs</i>	127
Figura 77. Pentagrama da melodia. A melodia é constituída por dez notas musicais distribuídas ao longo de dois compassos (padrão sonoro). De dois em dois compassos a melodia sobe um semi-tom e é repetida durante 12 vezes.	127
Figura 78. Média de respostas SIM e NÃO. O rácio entre as respostas SIM e Não é de 4:1.....	131
Figura 79. Interface de utilizador “FM”.	135
Figura 80. Representação geral de um circuito básico de síntese por FM aplicado ao Chromasom. A fm (f_0) é a frequência fundamental do Chromasom modulador e a fp (f_0) é a frequência fundamental do Chromasom portador. O índice de modulação (i) é a amplitude do Chromasom modulador e determina a quantidade de modulação exercida sobre cada uma das frequências das componentes do Chromasom portador.	136
Figura 81. Implementação em <i>Pure Data</i> de um circuito básico de síntese por FM aplicado ao Chromasom. O <i>patch</i> tem a particularidade de permitir introduzir e manipular, em tempo real os valores dos principais parâmetros da FM (i.e., fm, fp e o i), de forma a explorar e produzir diferentes timbres sonoros, que resultem musicalmente interessantes, para usar em composição musical digital. É possível observar os valores das frequências (Hz) e das amplitudes lineares de cada uma das componentes, quer do Chromasom modulador, quer do Chromasom portador. As amplitudes lineares são extraídas a partir da EEAAG e variam sempre, que se alteram os valores das frequências fundamentais (f_0), quer do Chromasom modulador, quer do Chromasom portador.	137
Figura 82. Interface de utilizador “chromasom_fm_presets”. Esta interface permite escutar os resultados sonoros obtidos a partir do circuito básico de síntese por FM aplicado ao Chromasom. É constituída por <i>presets</i> de sons percussivos e sons contínuos. Os <i>presets</i> encontram-se devidamente classificados e numerados na respectiva interface.	138
Figura 83. Exemplo de uma interface gráfica de utilizador directamente criada e editada na aplicação.	143
Figura 84. Interface de utilizador (MuDI versão 1.0). Em termos gerais, o Módulo Comunicação recebe, encaminha e trata a informação transmitida pela aplicação do dispositivo electrónico móvel. O Módulo Som está dedicado à geração sonora e à gravação áudio. O Módulo GUI permite visualizar a calibração do sistema, os pontos de referência e a relação entre os gestos realizados e o som gerado. O Módulo Vídeo permite o visionamento do filme e a reprodução musical.	146
Figura 85. Este diagrama de blocos apresenta (1) a estrutura geral da interface de utilizador, (2) as funções de cada um dos seus elementos e por último (3) o resultado final alcançado em termos de percepção audiovisual. No Módulo Comunicação os <i>subpatches</i> “pd sliders”, “pd buttons” e “pd accelerometer” estão inteiramente dedicados ao controlo dos geradores de som (Módulo Som). O <i>subpatch</i> “pd switches” é multi-funcional, no sentido, em que controla não só a geração do som, mas também a gravação da composição musical. Os <i>subpatches</i> “pd composition” e “pd video controls” estão unicamente destinados à reprodução da composição musical em sincronização com o filme seleccionado e com a paragem e reprodução das imagens do	

filme. Todos os <i>subpatches</i> , sem excepção estão ligados ao Módulo GUI para permitirem visualizar a manipulação dos controladores da aplicação do dispositivo electrónico móvel.	150
Figura 86. O diagrama de blocos fornece uma visão geral sobre a estrutura e o modo de operação do MuDI (versão 1.0). Através dele podemos ver o nome e o número total de controladores usados para tocar o MuDI. Foram escolhidos dois <i>sliders</i> , dois <i>buttons</i> , um acelerómetro e nove <i>switches</i> . Podemos ainda ver o tipo de gestos físicos usados para os manipular, a função atribuída a cada um deles e o resultado final obtido em termos de percepção audiovisual.	151
Figura 87. Representação do ponto de referência fiável do acelerómetro (ponto vermelho) e do dispositivo electrónico móvel colocado sobre uma mesa. A posição central do ponto de referência representa a calibração do sistema do MuDI. As suas coordenadas são 0° no eixo <i>y</i> e -90° no eixo <i>x</i> . Em termos de informática musical, as coordenadas correspondem a 0.5 de amplitude e à nota MIDI 63 de altura sonora.	152
Figura 88. Representação dos processos de recepção, encaminhamento e tratamento dos valores dos eixos <i>x</i> e <i>y</i> transmitidos pelo acelerómetro realizados no interior do <i>subpatch</i> “ <i>pd accelerometer</i> ” (Módulo Comunicação). A nota (a) representa as operações matemáticas e o processo de reajustamento realizados para obter os <i>y-values</i> (0 – 127, notas MIDI). A nota (b) representa as operações matemáticas e o processo de reajustamento realizados para se obterem os <i>x-values</i> (0 – 1, amplitude).	153
Figura 89. Representação dos pontos de referência (ponto vermelho) para a altura sonora e respectivos ângulos de rotação interna e externa do pulso realizados ao longo do eixo <i>y</i> ...	154
Figura 90. Representação dos pontos de referência (ponto vermelho) para a intensidade sonora e dos respectivos ângulos de flexão e extensão do pulso realizados ao longo do eixo <i>x</i>	154
Figura 91. Interface de utilizador <i>mudice\$1 - música de uma tecla só (a pretensiosa)</i> . 1- módulos de controlo: 1a) módulo controlo tecla “A”; 1b) módulo leitura melodia; 1c) módulo transposição Chromasom; 1d) tabela melodia; 1e) módulo transposição “pd noise”; 1f) módulo transposição geral; 1g) módulo volume geral. 2- módulos áudio: 2a) voz “pd Chromasom”; 2b) voz “pd som_puro”; 2c) voz “pd noise”; 2d) voz “pd osciladores”; 2e) voz “pd guit”; 2f) módulo volumes espacialização individuais; 2g) módulo reverberação. 3- módulo gravação.	158
Figura 92. Gerador de som <i>subpatch</i> “ <i>pd noise</i> ”.....	159
Figura 93. Gerador de som <i>subpatch</i> “ <i>pd osciladores</i> ”.....	160
Figura 94. Diagrama de blocos da interface de utilizador <i>mudice\$1 – música de uma tecla só (a pretensiosa)</i>	161
Figura 95. Melodia da <i>mudice\$1 – música de uma tecla só (a pretensiosa)</i>	162
Figura 96. Representação gráfica das envolventes das dinâmicas aplicadas às vozes da <i>mudice\$2 – a irreversível</i> . A linha contínua corresponde à envolvente aplicada à primeira voz (linha da vida), enquanto que a linha a tracejado diz respeito à envolvente aplicada à segunda voz (linha da morte). O conjunto das envolventes adquire a imagem de uma mesa de perfil, ou de um K deitado facialmente.	163
Figura 97. Interface de utilizador <i>mudice\$2 – a irreversível</i>	164
Figura 98. Diagrama de blocos da interface de utilizador <i>mudice\$2 – a irreversível</i>	165
Figura 99. Interface de utilizador <i>mudice\$3 - a irreversível</i> . A classificação número um, representa os módulos de controlo; a número dois, os módulos áudio e a número três, o módulo de gravação.	167
Figura 100. Diagrama de blocos da interface de utilizador <i>mudice\$3 – a perpétua</i>	168
Figura 101. Interface de utilizador <i>Perpetual Melody – contrasting moments</i>	170
Figura 102. Diagrama de blocos da interface de utilizador <i>perpetual melody – contrasting moments</i>	171
Figura 103. Interface de utilizador <i>sonificação - filmes para música\$1</i> (MuDI versão 1.0).	174
Figura 104. Interface de utilizador <i>sonification - films for music\$2</i> (MuDI versão 1.1).	174

ÍNDICE DE TABELAS E GRÁFICOS

Tabela 1. Relação entre as alturas sonoras, os comprimentos de onda dos sons e a dimensão da cabeça e do ouvido externo na audição. Adap. de Mathews (1999, p. 3).	14
Tabela 2. Apresentação das frequências individuais dos Sons de Shepard, que se encontram representadas nas listas da Figura 39. As linhas (s) correspondem à variável (t) da Equação 2.	83
Tabela 3. Notas musicais das frequências individuais obtidas a partir do procedimento representado na Figura 40.	84
Tabela 4. Amplitudes individuais das componentes.	86
Tabela 5. Frequências individuais do Chromasom.	103
Tabela 6. Notas musicais das componentes do Chromasom.	103
Tabela 7. Valores das amplitudes individuais do Chromasom obtidos a partir da Equação 10. Verifica-se que as amplitudes individuais aumentam gradualmente desde a primeira até à sétima componente e diminuem desde à sétima até à nona componente de cada um dos sons.	120
Tabela 8. Apresentação dos resultados obtidos no exercício número um.	129
Tabela 9. Apresentação dos resultados obtidos no exercício número dois.	130
Tabela 10. Pontos de referência do eixo y, coordenadas dos movimentos de rotação interna e externa do pulso e nota MIDI correspondente obtida para a altura sonora.	155
Tabela 11. Pontos de referência do eixo x, coordenadas dos movimentos de flexão e extensão do pulso e o valor correspondente obtido para a intensidade sonora.	155

INTRODUÇÃO

No início do nosso percurso de compositor começámos por compor obras musicais acusmáticas, registadas em suporte fixo e difundidas através de orquestras de altifalantes. Na composição deste tipo de obras, utilizávamos normalmente como fontes sonoras sintetizadores MIDI e gravações directas de sons provenientes do meio ambiente que nos rodeava (e.g., gravações ao vivo de sessões de cinema, concertos, etc.), ou sons obtidos através da manipulação de objectos do nosso dia-a-dia (e.g., subir e descer escadas de garagens de prédios, abrir e fechar portas, tocar guitarra com vários objectos, etc.).

Apesar de procurarmos que os sons fossem sequenciados e processados de maneira a que o ouvinte não identificasse a fonte sonora, o facto é que, para nós, essa identificação estava sempre presente. Por outro lado, sentíamo-nos condicionados, por não termos um controlo efectivo sobre estas mesmas fontes sonoras, no que respeita, não só a questões de qualidade sonora ou musical, mas também, por não conseguirmos eliminar totalmente a identificação da sua origem.

Assim, para conseguirmos um maior controlo, não só sobre as fontes sonoras, mas também sobre os sons, e para aumentar o nível de abstração sonoro-musical, no que respeita à identificação das fontes sonoras decidimos dirigir a nossa atenção para a síntese de som digital. Neste novo paradigma, a construção de fontes sonoras a partir de osciladores, conferiu-nos a possibilidade de obtermos um maior controlo, não só sobre as fontes sonoras, mas também sobre os sons. No entanto, no que respeita aos níveis de abstração sonoro-musical sentíamos que poderíamos ir ainda um pouco mais além, e que deveríamos procurar uma fonte sonora que possuísse características psicoacústicas únicas em termos timbricos e de percepção auditiva.

Com estes pressupostos, fomos procurar a nossa resposta ao universo das ilusões sonoras. Foi então, que nos deparamos com a Ilusão Sonora da Escala de Shepard e com o Som de Shepard, que representam casos peculiares de percepção auditiva, no que respeita,

à percepção de altura sonora. Não só o efeito sonoro produzido pela Ilusão Sonora da Escala de Shepard nos causou um grande fascínio, como o Som de Shepard nos permitiu vislumbrar, que estaríamos possivelmente perante um fonte sonora que nos poderia oferecer um amplo potencial musical por explorar. No entanto, paralelamente tivemos consciência, que não bastava somente reproduzir o Som de Shepard e usá-lo em projectos musicais, mas teríamos de fazer um pouco mais, por exemplo, melhorá-lo em termos tímbricos, aplicar-lhe estratégias de síntese sonora inovadoras, através da utilização de uma envolvente espectral não simétrica, submetê-lo a experiências auditivo-musicais mais avançadas, ou seja, aplicá-lo a melodias constituídas por saltos intervalares não regulares e paralelamente utilizá-lo como gerador de som em projectos musicais.

Com efeito, as ilusões sonoras sempre geraram um enorme fascínio sobre os seres humanos, sendo normalmente originadas (na presença de estímulos) a partir de fenómenos físicos ou naturais, como o eco ou as vozes que se ouvem no som de água a correr, no sopro do vento, ou em outras fontes sonoras quaisquer.

No entanto, só na segunda metade do século XX, fruto da grande precisão de síntese de som gerada através de computador foi possível criar ilusões sonoras a partir de processos artificiais e em contexto de laboratório (Gregory 1997, p. 1122).

Em 1964, nos Laboratórios da Empresa Telefónica Bell, em New Jersey (EUA), Shepard sintetizou em computador uma escala musical dividida em 12 partes iguais, que ao ser escutada, cria a sensação de perpetuidade sonora, como acontece no plano visual com a Escada Interminável de Escher (1960), ou com as linhas oblíquas de um Poste de Barbeiro em movimento. Esta peculiar ilusão dá pelo nome de Ilusão Sonora da Escala de Shepard. Esta ilusão, relacionada com o domínio das frequências deu origem a algumas variantes, por exemplo o Glissando Perpétuo de Risset (1969) ou o Paradoxo do Trítone de Deutsch (1986).

Decidimos estudar a Ilusão Sonora da Escala de Shepard, por ser aquela que mais nos fascinou em termos sonoros e por ter revelado um grande potencial musical por explorar. Parte deste potencial, reside no facto, da ilusão se apresentar sob a forma de uma escala musical, e simultaneamente, por ser constituída por sons (i.e., o Som de Shepard), que apresentam a peculiaridade de serem ambíguos, em termos de percepção de altura sonora, permitindo, desta maneira, criar novas possibilidades de geração de som e de escuta musical.

Através de processos de síntese de som digital é possível isolar os atributos auditivos (normalmente correlacionados) que se encontram relacionados com a percepção de altura sonora de sons complexos (sons constituídos por mais do que uma componente) (Fugiel 2011, p. 415), como é o caso do ruído e do Som de Shepard. Os atributos auditivos são: a altura espectral (*tone height*) e a altura tonal (*tone chroma*). A altura espectral representa a dimensão vertical dos sons, enquanto que a altura tonal representa a sua dimensão horizontal (Deutsch 2008, p. 589).

Assim, esta dissertação propõe a criação de uma fonte sonora baseada no Som de Shepard com o objectivo de ser utilizada em projectos musicais. Consequentemente, criámos um método de gerar o Som de Shepard através de uma envolvente espectral alternativa, tanto quanto sabemos, nunca antes usada para reproduzir a Ilusão Sonora da Escala de Shepard, mesmo quando esta, é aplicada a uma melodia constituída por uma sequência de saltos intervalares não regulares. A melodia foi intitulada de Ilusão Sonora da Melodia Perpétua, porque quando é escutada, cria a sensação de perpetuidade sonora, como acontece na Ilusão Sonora da Escala de Shepard. Para além disso, utilizamos a fonte sonora proposta como gerador de som na concepção de um instrumento digital multimédia e em composição musical.

No que respeita à composição musical, utilizando a fonte sonora como gerador de som, iniciámos dois ciclos de obras musicais designados por (1) Música Digital Contemporânea Erudita (MuDiCE), onde a síntese de som e a sequenciação musical foram realizadas integralmente em *Pure Data*¹ e (2) Sonificação – filmes para música, cujas composições, realizadas em tempo real, foram obtidas através de um instrumento digital multimédia, por nós concebido e tocado.

Principais Contribuições

Vários autores estudaram, quer o Som de Shepard, quer a Ilusão Sonora da Escala de Shepard. Nesses estudos, utilizaram-se envolventes espectrais com configuração simétrica (Risset, 1977; Pollack, 1978; Burns, 1981; Deutsch et al., 1986; Ueda e Ohgushi, 1987; Nakajima *et al.*, 1988). Um dos contributos desta dissertação, passa pelo facto de, na revisitação que fazemos do Som de Shepard e ao contrário dos autores

¹ O Pure Data é uma linguagem de programação gráfica. Mais informação em <http://puredata.info/>

anteriormente mencionados, utilizarmos uma envolvente espectral alternativa de configuração não simétrica, baseada na curva de igual percepção de intensidade sonora de 40 *phons* de Fletcher e Munson (1933). Outro contributo, consiste na exploração das consequências da utilização da respectiva ilusão sonora aplicada a uma melodia constituída por uma sequência de saltos intervalares não regulares (Figura 77). Desta exploração resultou a criação da Ilusão Sonora da Melodia Perpétua (2012).

Risset (1938), na qualidade de compositor utilizou directamente a versão original da Ilusão Sonora da Escala de Shepard, tocada de forma descendente no segundo movimento (Fall) da sua obra musical intitulada “*Computer Suite for Little Boy* (1968)”. No que respeita à composição musical, o contributo desta dissertação, passa por expandir a utilização musical da Ilusão Sonora da Escala de Shepard, realizada por um lado, através da aplicação da ilusão à melodia representada na Figura 77 (Ilusão Sonora da Melodia Perpétua) e por outro lado, na utilização da respectiva melodia como conteúdo musical. Ambas as abordagens foram materializadas através da obra musical intitulada “*Perpetual Melody – contrasting moments* (2012)”.

Durante a fase de pesquisa, ao efectuarmos uma procura na Internet deparámo-nos com várias demonstrações audiovisuais da ilusão, que por vezes apresentam variantes em relação à descrição de Shepard, sob o ponto de vista da construção da ilusão e do seu mecanismo de funcionamento. Nas demonstrações encontradas na Internet (em sítios não oficiais)², as componentes do Som de Shepard encontram-se em movimento constante por ordem de entrada, do registo grave para o registo agudo do espectro sonoro. Para além disso, as componentes encontram-se envolvidas por uma única envolvente espectral fixa. Nesta dissertação, as componentes são fixas e cada som possui uma envolvente espectral, que apesar de fixa é simultaneamente ajustada a cada um dos sons. O Som de Shepard é um som estático e ocupa sempre a mesma posição ao longo da escala. Há, no entanto, um movimento, mas é um movimento aparente que resulta não da construção física ou arquitectónica da ilusão, mas da sua percepção auditiva. É por esta razão que a Ilusão Sonora da Escala de Shepard nos revela que o fenómeno de movimento aparente também existe na audição (Shepard 1999, p. 126). Desta maneira, com a abordagem que fazemos da respectiva ilusão estamos também a contribuir para um melhor esclarecimento, no que respeita à sua construção e mecanismo de funcionamento.

² Exemplos em <http://www.youtube.com/watch?v=ev9hrqkhWsM>;
<http://www.cs.ubc.ca/nest/imager/contributions/flinn/Illusions/ST/st.html>;
<http://www.netalive.org/tinkering/shepard-effect/>

Pelo facto, desta dissertação ser um estudo sobre a forma como percebemos os sons, como nos confrontamos com as ilusões e com os sons que não compreendemos, estamos também a contribuir para conhecer e compreender melhor os mecanismos de funcionamento da percepção auditiva. Para além disso, pelo facto da fonte sonora proposta ser um som ambíguo em termos de percepção de altura sonora estamos também a ampliar as nossas possibilidades de escuta sonora e musical.

Aplicações da Fonte Sonora

Através da fonte sonora proposta iniciámos um ciclo de obras musicais chamado Música Digital Contemporânea Erudita (MuDiCE), do qual destacamos:³

- *mudice\$1 – música de uma tecla só (a pretensiosa)*, por estrear;
- *mudice\$2 – a irreversível*, estreada no concerto de comemoração dos 25 anos da Miso Music Portugal em Abril de 2010;
- *mudice\$3 – a perpétua*, estreada no Festival Música Viva em Setembro de 2010;
- *Perpetual Melody – contrasting moments*, por estrear.

Paralelamente, no contexto da interpretação musical com dispositivos electrónicos móveis (*mobile music performance*) e da composição musical em tempo real, concebemos um instrumento multimédia para compor e tocar música para filmes, que utiliza para além de outros geradores de som, a fonte sonora anteriormente mencionada. O instrumento multimédia chama-se MuDI. O MuDI permite compor e tocar música, gravar o som da interpretação, obter uma partitura visual, acompanhar não só as imagens do filme, mas também a relação entre os gestos realizados e o som gerado, obter um controlo efectivo sobre o som e consequentemente alcançar grande expressão musical.

Através do MuDI, iniciámos um novo ciclo de obras musicais com o nome de *sonificação – filmes para música*.⁴ Deste ciclo, fazem parte duas composições. A primeira, *sonificação – filmes para música\$1*, foi estreada em Setembro de 2009, no 12.º

³ As composições musicais podem ser consultadas através do DVD em anexo, no directório MuDiCE. A designação \$1, deve ler-se variável um; \$2, variável dois, e assim sucessivamente.

⁴ As composições musicais, os excertos vídeo dos concertos e as apresentações relacionadas com o MuDI podem ser consultados no DVD em anexo, no directório MuDI.

Simpósio Brasileiro de Computação Musical (SBCM). A segunda, *sonification – films for music*§2 foi apresentada, em versão curta em Outubro de 2011, no 2.º Fórum Internacional Itinerários Musicais - Música e Gesto, UNL, Portugal, e ainda na 12.ª Conferência Internacional *New Interfaces for Musical Expression* (NIME), Universidade de Michigan, Ann Arbor – Michigan, EUA, Maio de 2012.

Estrutura da Dissertação

Em termos gerais, a dissertação encontra-se dividida em três partes. Na Parte I (Termos, Noções e Conceitos) abordamos os termos, as noções e os conceitos relacionados com o sistema auditivo, a psicoacústica, a percepção e as ilusões. Estes têm como finalidade, para além da revisão bibliográfica, concentrar aqui alguns conceitos básicos, que podem funcionar como uma breve e genérica introdução aos mesmos, e ainda, estabelecer um fio condutor narrativo entre as problemáticas tratadas no início desta dissertação, relacionadas com o som e a citação da escritora Maria Gabriela Llansol (1931-2008), que se encontra no final das Conclusões e Direcções Futuras do nosso trabalho.

Esta abordagem permitirá também contextualizar e compreender determinadas questões relacionadas com fenómenos auditivos como, por exemplo, os batimentos mencionados na descrição da obra musical *mudice*§3 – *a perpétua*, o efeito de *Doppler* (que poderá surgir quando movimentamos a cabeça em frente às colunas que difundem o som do material multimédia - obras musicais, *patches* e ficheiros áudio) ou os princípios gestálticos de agrupamento (que nos ajudam a compreender o mecanismo de funcionamento das ilusões sonoras em geral e, em particular, das Ilusões Sonoras da Escala de Shepard e da Melodia Perpétua).

No que respeita ao sistema auditivo, para além da sua estrutura e funcionamento abordamos sobretudo, os aspectos considerados mais pertinentes, no sentido, em que condicionam a maneira como percebemos os sons, a música e alguns fenómenos psicoacústicos.

Em relação à psicoacústica apresentamos o percurso dos processos entre os eventos físicos que ocorrem no mundo que nos rodeia e a representação que fazemos desses mesmos eventos. Apresentamos também uma analogia sobre o principal “problema”, que o nosso sistema auditivo enfrenta ao usar o tímpano como porta de

entrada para a audição. Realizamos um teste para verificar se existe uma relação directa entre a percepção auditiva de dois intervalos musicais (e.g., entre um intervalo de quinta perfeita e um intervalo de segunda menor) e a sua representação gráfica em forma de onda sonora. Por último, abordamos o fenómeno dos batimentos entre os sons e a sua relação com a percepção de dissonância.

Sobre a percepção abordamos as principais teorias (Teoria directa de Gibson, 1966 vs. Teoria construtivista de Gregory, 1970) e os principais mecanismos (e.g., inferência inconsciente, percepção de preenchimento e princípios gestálticos de agrupamento).

No que respeita às ilusões falamos sobre a sua definição e classificação, até chegarmos ao conceito de ilusão cognitiva. Realizamos um breve estudo sobre os processos de geração das ilusões e falamos ainda, sobre os mecanismos de actuação das ilusões na percepção humana.

Na Parte II (Ilusão Sonora da Escala de Shepard) abordamos a respectiva ilusão, no que respeita, à sua analogia com a Ilusão Visual da Escada Interminável de Escher e com as linhas oblíquas de um Poste de Barbeiro em movimento. Para além disso, falamos sobre (1) a noção de ciclo estranho segundo Hofstadter (1979), (2) o Som e a Escala de Shepard, (3) a representação bidimensional da percepção de altura de sons complexos, cujas componentes se encontram separadas por intervalos de oitava, (4) a percepção auditiva da Ilusão Sonora da Escala de Shepard e os factores que contribuem para a sua realização, e ainda (5) o Som de Shepard e a Ilusão Sonora da Escala de Shepard revisitados por vários autores. Para além destas considerações de carácter mais teórico, apresentamos as etapas que constituem o processo de implementação da respectiva ilusão sonora em síntese de som digital. Referimo-nos concretamente à sua reconstituição e reprodução áudio. Com o propósito de verificar se não ocorreu nenhum erro durante o processo de reconstituição e reprodução áudio da respectiva ilusão realizamos uma análise espectral ao Som de Shepard, através de uma transformada rápida de Fourier real (RFFT). Não ficando totalmente satisfeitos com o resultado sonoro obtido na reprodução áudio da ilusão decidimos construir a nossa própria fonte sonora.

Assim, na Parte III (Desenvolvimento e Aplicações) propomos a construção de uma fonte sonora baseada no Som de Shepard, com o objectivo de ser utilizada como gerador de som em projectos musicais. Concebemos um método para gerar o Som Shepard, através de uma envolvente espectral alternativa, baseada na curva de igual percepção de intensidade sonora de 40 *phons* de Fletcher e Munson (1933). Na continuidade desta parte da dissertação, submetemos a fonte sonora a duas séries de

exercícios musicais e a uma experiência auditivo-musical, com o objectivo de verificar, se o efeito da Ilusão Sonora da Escala de Shepard, isto é, se a sensação de perpetuidade sonora se mantém durante a repetição dos exercícios musicais e da respectiva experiência. Seguidamente, aplicamos à fonte sonora um circuito básico de síntese por Modulação de Frequência (FM) e por último, falamos da utilização da fonte sonora como gerador som, quer na concepção de um instrumento digital multimédia (MuDI), quer em composição musical.

No que concerne à composição musical, apresentamos dois ciclos constituídos por obras musicais compostas integralmente durante o período em que decorreu esta investigação. Em ambos os ciclos (1) a fonte sonora desempenha um papel preponderante na geração de som, (2) a aplicação *Pure Data* é usada como ambiente exclusivo de programação, no que respeita, à síntese de som e à sequenciação musical, (3) o computador portátil é utilizado como auxiliar de composição, e simultaneamente, como instrumento musical e por último (4) um dispositivo electrónico móvel comum é utilizado como controlador musical.

Nas conclusões e direcções futuras apresentamos as ideias fundamentais que pretendemos realizar e desenvolver na continuidade deste trabalho de investigação. Fazem ainda parte da estrutura da dissertação o anexo número um, que contém a lista dos artigos publicados e as conferências realizadas durante o doutoramento em eventos com escrutínio científico; o anexo número dois, relacionado com a produção artística realizada também durante o doutoramento, onde se apresentam as listas das obras musicais compostas e estreadas, os concertos realizados, a publicação em CD e a referência à divulgação radiofónica de uma das obras; o anexo número três, onde se faz a listagem do conteúdo do DVD que acompanha a dissertação e o anexo número quatro, que contém o formulário aplicado no âmbito do estudo estatístico realizado.

Gostaríamos de fazer notar, que na dissertação optámos por empregar a notação anglo-saxónica para as alturas dos sons. Por exemplo, o Lá 4, de 440 Hz é notado como A4.

Por motivos de programação informática e simultaneamente, para que a nomenclatura numérica, que aparece no texto escrito seja coerente com a das Figuras e das Tabelas apresentadas, optámos também, por usar a nomenclatura numérica anglo-saxónica, em que o ponto representa as unidades (e.g., 4.083 Hz) e os milhares são representados da seguinte maneira: 4083 Hz.

A maioria dos termos encontrados na literatura estrangeira, foram traduzidos e adequados à língua portuguesa. Por exemplo, o termo *Pitch*, foi traduzido para percepção de altura sonora. Desta maneira, pensamos estar a contribuir para o enriquecimento e expansão lexical da nossa língua materna. Mais, vide nesta dissertação, o glossário abreviado de termos, que se encontra logo a seguir à bibliografia.

O material multimédia (áudio, vídeo e *patches* de *Pure Data*), mencionado nas notas de rodapé, encontra-se classificado e está disponível no DVD que acompanha a dissertação.

Os *patches* realizados em *Pure Data* possuem instruções de funcionamento. No entanto, os utilizadores que não se sentirem familiarizados com a aplicação *Pure Data*, têm, como alternativa, a possibilidade de aceder aos mesmos conteúdos audiovisuais através da reprodução dos ficheiros áudio e vídeo correspondentes.

PARTE I – TERMOS, NOÇÕES E CONCEITOS

Há uma sequência de processos a percorrer entre os eventos físicos que ocorrem no mundo que nos rodeia e a representação que fazemos desses mesmos eventos. Esses processos incluem (1) a origem da energia através de objectos ou eventos externos, (2) a transmissão da energia através do espaço entre o evento e o observador, (3) a recepção e o processamento da energia pelos receptores sensoriais do observador e (4) a transmissão dos sinais para o cérebro.

Na audição, podemos ordenar e esquematizar estes processos da seguinte maneira: (1) origem da energia através de fontes sonoras (e.g., vozes humanas), (2) propagação das ondas sonoras através de um meio físico (e.g., ar) e por último (3) a transdução do sinal físico em electroquímico, que ocorre ao nível do ouvido e do cérebro (Emanuel, Maroonroge e Letowski 2009, p. 307; Plack 2004, pp. 2-6; Mathews 1999, p.1; Pierce 1999, p. 89).

Podemos generalizar, que o ouvido humano realiza uma transdução de energia semelhante à de um microfone (i.e., converte um tipo de energia numa energia de outro tipo). O microfone capta o sinal físico (as variações de pressão de ar) e reconstrói-o enviando impulsos eléctricos para um gravador áudio. Os nossos ouvidos comportam-se como uma espécie de microfones naturais, porque captam e transformam o sinal físico em estímulos electroquímicos, que ao serem enviados para o cérebro, permitem criar uma representação mental do sinal físico recebido. Em ambos os casos, estamos perante um processo de transdução de sinal, no qual os receptores eléctricos (no caso do microfone) e os receptores sensoriais (no caso da audição) funcionam como transdutores.

A energia é submetida a várias metamorfoses, desde a origem até à sua percepção. Referimo-nos concretamente à transformação da energia molecular (i.e., o ar) em energia eléctrica (i.e., em impulsos eléctricos) e esta, por sua vez, em energia cerebral (i.e., impulsos nervosos).

Entendemos que esta energia só se transforma realmente em som quando é efectivamente interpretada pelo ouvinte. Caso contrário, podemos estar perante uma espécie de iliteracia sonora equivalente à iliteracia das palavras (não basta conhecer as palavras ou saber ler, é necessário saber interpretá-las para sabermos o que nos querem dizer).

De seguida, iremos abordar o sistema auditivo, no que respeita à sua estrutura anatómica, começando pelo ouvido e terminando no cérebro.

I SISTEMA AUDITIVO

Vamos realizar uma abordagem sintética sobre o sistema auditivo, incidindo sobretudo, nos aspectos pertinentes, que determinam a maneira como percebemos os sons, a música e alguns fenómenos psicoacústicos.

I.1 Ouvido

A estrutura do ouvido encontra-se dividida em três partes (ouvido externo, médio e interno). Apesar de terem características e funções diferentes, todas elas se relacionam entre si (Figura 1).

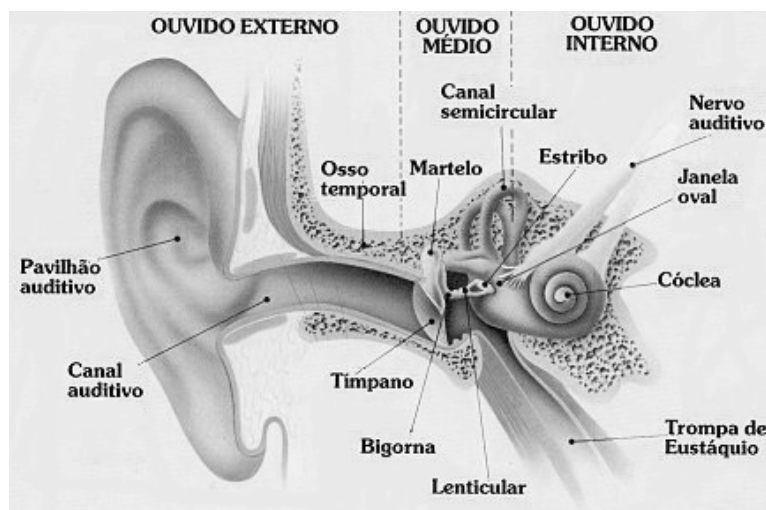


Figura 1. Representação das estruturas do ouvido. As estruturas e os seus elementos não se encontram representados à escala real. Adap. de Mathews (1999, p. 2).

I.1.1 Ouvido Externo

O ouvido externo funciona como captador, amplificador e filtro direccional dos sons. Recolhe e encaminha as moléculas de ar ao longo do canal auditivo fazendo vibrar o tímpano. É constituído pelo pavilhão e canal auditivos. O canal auditivo prolonga-se até ao tímpano e possui mais ou menos 2.54 cm de comprimento e 0.6 cm de diâmetro.

O pavilhão auditivo, a cabeça e os ombros podem interferir na captação das ondas sonoras produzindo determinados efeitos. Os efeitos, tornam-se significativos sobretudo, quando o comprimento de onda, das ondas sonoras é proporcional à dimensão destas mesmas estruturas. Por exemplo, os sons de alta frequência, como possuem comprimentos de onda muito pequenos são reflectidos pelos ombros em direcção ao ouvido externo (Pierce 1999, p. 98). De acordo com Mathews (1999, p. 2) a influência dos ombros e da cabeça sobre os sons faz-se sentir em alturas sonoras superiores a 1760 Hz (i.e., A6) e a interferência do ouvido externo em sons acima dos 7040 Hz (i.e., A8) (Tabela 1).

Tabela 1. Relação entre as alturas sonoras, os comprimentos de onda dos sons e a dimensão da cabeça e do ouvido externo na audição. Adap. de Mathews (1999, p. 3).

Altura do som	Frequência (Hz)	Comprimento de onda (m)	Cabeça (m)	Ouvido externo (m)
A2	110	3.124		
A3	220	1.574		
A4	440	0.787		
A5	880	0.381		
A6	1760	0.195	0.195	
A7	3520	0.096		
A8	7040	0.048		0.048

De acordo com a Tabela 1, se se considerar unicamente o som directo (i.e., a parte das ondas sonoras que chega directa e ininterruptamente aos ouvidos, desde a fonte sonora até à posição do ouvinte) uma frequência de 1760 Hz ao atingir o ouvido, que se encontra mais próximo da fonte sonora, como não possui um comprimento de onda superior à largura da cabeça do ouvinte (0.195 m) não conseguirá contorná-la e atingir o ouvido mais afastado, acabando por ser absorvida.

No caso da música, a lateralização extrema que as frequências superiores a 7040 Hz poderia implicar, fica atenuada, porque segundo Mathews (1999, p. 2) nenhum instrumento musical produz normalmente uma frequência fundamental com essa altura.

A influência da cabeça também é preponderante na percepção da perspectiva sonora (i.e., na localização espacial das fontes sonoras). Segundo Mathews (1999, p. 3) a cabeça, ao criar uma “sombra sonora” sobre o ouvido externo, desempenha um papel preponderante na determinação da localização lateral do som, ou seja, se o som vem do lado esquerdo ou do lado direito do ouvinte. Para além disso, interfere também no tempo de chegada dos estímulos auditivos a ambos os ouvidos. Diferenças entre os 20 e os 40 milissegundos podem produzir alterações apreciáveis na percepção espacial. Este fenómeno é conhecido por Efeito de Haas ou Efeito de Precedência (Pierce 1999, p. 92). É um fenómeno psicoacústico que nos permite localizar a proveniência dos sons (Warren 2008, p. 113). Acontece, devido à geometria da cabeça humana que funciona como uma barreira intermédia entre os dois ouvidos. O som directo atinge primeiro o ouvido mais próximo da fonte sonora e em seguida o ouvido mais afastado. O Efeito de Haas ou Efeito de Precedência refere-se também ao primeiro estímulo sonoro que chega à posição do ouvinte tornando-se este precedente em relação aos estímulos seguintes. Caso os outros estímulos sonoros não cheguem dentro do intervalo de tempo acima mencionado interpretá-los-emos como estímulos distintos.

De acordo com Mathews (1999, p. 4) o pavilhão auditivo desempenha um papel importante na absorção da intensidade sonora e na determinação da localização dos sons, ou seja, se estes chegam pela frente ou por trás do ouvinte. A posição anatómica frontal do pavilhão auditivo favorece a recepção de sons que chegam pela frente (Emanuel, Maroonroge e Letowski 2009, p. 308).

Existem também outras pistas auditivas que nos ajudam a determinar a localização ou a direcção das fontes sonoras. São elas, a diferença de intensidade e a diferença temporal interaurais. A diferença de intensidade interaural é a diferença de intensidade que as ondas sonoras provocam ao atingir ambos os ouvidos (Cheng e Wakefield 2001, p. 231). Está relacionada com os efeitos de absorção e reflexão gerados pelo corpo e pelas estruturas da cabeça. Afecta, sobretudo, o ouvido que se encontra mais afastado da fonte sonora (Emanuel, Maroonroge e Letowski 2009, p. 308). Por exemplo, o som que chega pelo lado esquerdo da nossa cabeça percorre uma distância mais curta para atingir o ouvido esquerdo do que para atingir o ouvido mais afastado (i.e., o ouvido direito). Por

isso, sentimos o estímulo sonoro mais intenso no ouvido esquerdo e consequentemente como sendo proveniente desse mesmo lado (Scharine, Maroonroge e Letowski 2009, pp. 457-458; Warren 2008, p. 54; Mathews 1999, p. 16; Pierce 1999, p. 81).

A diferença temporal interaural é a diferença sentida no tempo de chegada das ondas sonoras a ambos os ouvidos (Cheng e Wakefield 2001, p. 231). Por exemplo, o som que chega pelo lado direito da nossa cabeça viaja menos tempo para atingir o ouvido direito do que para atingir o ouvido mais afastado (i.e., o ouvido esquerdo). Por esse motivo, sentimo-lo como sendo proveniente do lado direito (Emanuel, Maroonroge e Letowski 2009, p. 308; Warren 2008, p. 54; Mathews 1999, p. 16; Pierce 1999, p. 81).

A diferença de intensidade interaural é eficaz na localização de sons de alta frequência, enquanto que a diferença temporal interaural é mais eficaz na localização das baixas frequências (Emanuel, Maroonroge e Letowski 2009, p. 308; Plack 2004, p. 17). Como a dimensão da cabeça é maior do que os comprimentos de onda das altas frequências, as ondas sonoras ao não serem difractadas (i.e., ao não contornarem a cabeça) fazem com que a diferença de intensidade sentida entre ambos ouvidos seja significativa. Por outro lado, como as baixas frequências possuem comprimentos de onda superiores à dimensão da cabeça, estas, ao serem difractadas fazem com a diferença temporal, na chegada das ondas sonoras a ambos os ouvidos seja relevante (Scharine, Maroonroge e Letowski 2009, pp. 457). Portanto, o ouvido mais próximo da fonte sonora receberá o som mais cedo e com maior intensidade (Emanuel, Maroonroge e Letowski 2009, p. 308).

Todos estes factores auditivos, anteriormente mencionados são importantes na localização dos sons no espaço, como atesta o processo de medição designado por *Head Related Transfer Function* (HRTF). Neste processo, a localização de uma fonte sonora é determinada a partir da diferença entre as ondas sonoras que atingem os ombros, a cabeça, o pavilhão auditivo e as ondas sonoras absorvidas pelo canal auditivo (Emanuel, Maroonroge e Letowski 2009, p. 308; Warren 2008, p. 54; Cheng e Wakefield 2001, p. 231).

Este processo tem sido utilizado, não só nas áreas de investigação da acústica e da psicoacústica, mas também, na arte e no entretenimento, por exemplo, através da simulação de espaços sonoros virtuais tridimensionais (Duda 1993, p. 1).

I.1.2 Ouvido Médio

O ouvido médio é a cavidade de ar que fica por trás do tímpano. Esse espaço também é conhecido por caixa timpânica. Digamos que o tímpano estabelece a fronteira entre o ouvido externo e o ouvido médio.

O tímpano vibra de acordo com a frequência das ondas sonoras. É no ouvido médio que se encontram alojados os ossos vestibulares (martelo, bigorna e estribo) (Figura 2).

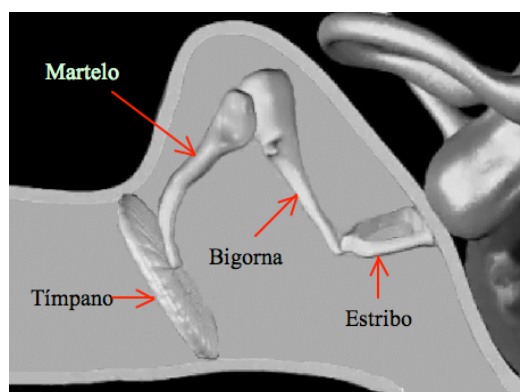


Figura 2. Principais estruturas do ouvido médio. Adap. de Warren (2008, p. 7).

Como os ossos vestibulares se encontram ligados através do martelo ao tímpano vibram com a mesma frequência do tímpano, e simultaneamente, amplificam a sua vibração. A intensidade de vibração da bigorna é aproximadamente 15 vezes superior à do tímpano. Esta característica, confere-nos a possibilidade de ouvirmos os sons menos intensos. O último osso da cadeia vestibular, o estribo está ligado a uma fina membrana existente na extremidade basal da cóclea que se denomina janela oval (Figura 3).

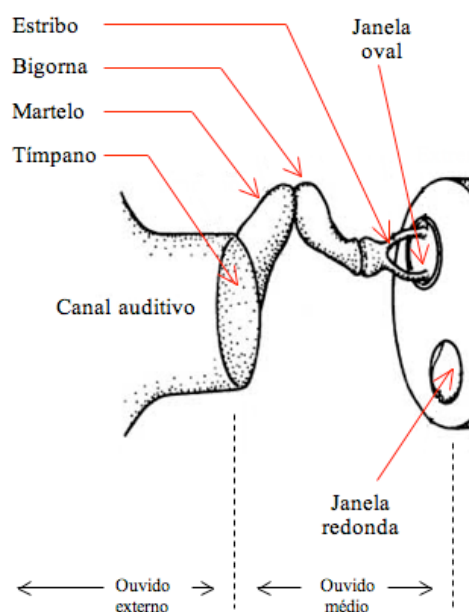


Figura 3. Ilustração da cadeia dos ossos vestibulares e da ligação do estribo à janela oval. Adap. de Warren (2008, p. 8).

Os músculos do ouvido médio previnem o excesso de movimento dos ossos vestibulares protegendo o ouvido interno de sons muito intensos. Este fenómeno é conhecido por Reflexo Acústico ou Auditivo (Emanuel, Maroonroge e Letowski 2009, p. 316). O reflexo, consiste na contracção muscular involuntária, que ocorre no ouvido médio, em resposta a estímulos sonoros de alta intensidade. Os músculos responsáveis por este fenómeno são: o músculo *stapedius* e o músculo tensor timpânico. Quando somos confrontados com sons intensos (e.g., ruídos de motores, vozes, etc.), os músculos contraem-se. O *stapedius* puxa o estribo afastando-o da janela oval e o tensor timpânico afasta o martelo do contacto com o tímpano. Este reflexo diminui a energia vibracional transmitida para o interior da cóclea protegendo, desta maneira, o ouvido interno (Warren 2008, p. 8). O reflexo, reduz os níveis de pressão sonora, que alcançam o ouvido interno aproximadamente em 20 dB. Contudo, como são músculos que se fatigam muito rapidamente vão perdendo a sua eficácia à medida que nos expomos por longos períodos a sons muito intensos.

I.1.3 Ouvido Interno

No ouvido interno encontramos a cóclea, os canais semicirculares e o nervo

auditivo. O fluido e as células nervosas dos canais semicirculares não têm função na audição. Desempenham unicamente funções na detecção de movimentos e na manutenção do equilíbrio do corpo.

A cóclea tem a forma de caracol ou de um tubo enrolado em espiral. Quando esticada pode atingir aproximadamente três centímetros e meio de comprimento (Warren 2008, p. 9). Tem como função, converter a energia mecânica gerada pelos movimentos dos ossículos vestibulares em impulsos electroquímicos, que irão ser transmitidos pelo sistema nervoso central até aos centros auditivos do cérebro (Emanuel, Maroonroge e Letowski 2009, p. 317; Plack 2004, pp. 5-6; Mathews 1999, p. 1; Pierce 1999, p. 89). Encontra-se dividida, em toda a sua extensão por três secções preenchidas por líquido (Figura 4).

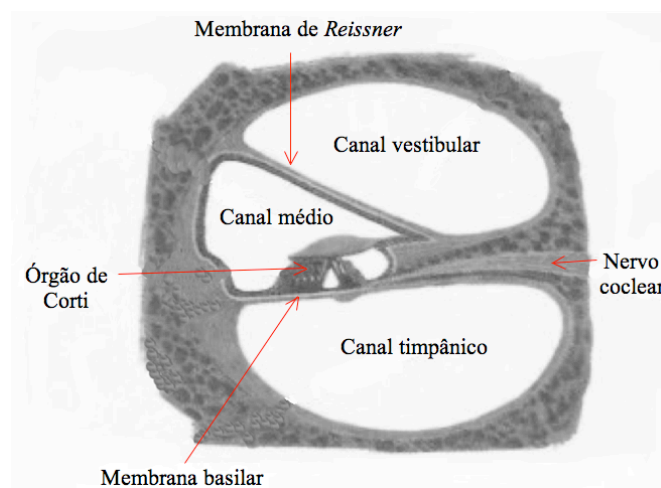


Figura 4. Corte longitudinal da cóclea. Adap. de Mathews (1999, p. 6).

A primeira secção denomina-se canal vestibular (*scala vestibuli*) e está ligada à janela oval. A última secção, chama-se *scala tympani* (canal timpânico) e encontra-se ligada à janela redonda. Estas duas secções, unem-se apenas no fim da cóclea através da *helicotrema*. Encontram-se ainda, separadas por uma terceira secção denominada de canal médio (*scala media*) ou ducto coclear (Warren 2008, p. 9).

A separação entre as três secções referidas é efectuada por duas membranas. A separar o canal vestibular do canal médio encontra-se a membrana de *Reissner* e entre o canal médio e o canal timpânico encontra-se a membrana basilar. A membrana basilar é muito importante, pois é ela que suporta o órgão de Corti (órgão receptor da audição) (Figura 5).

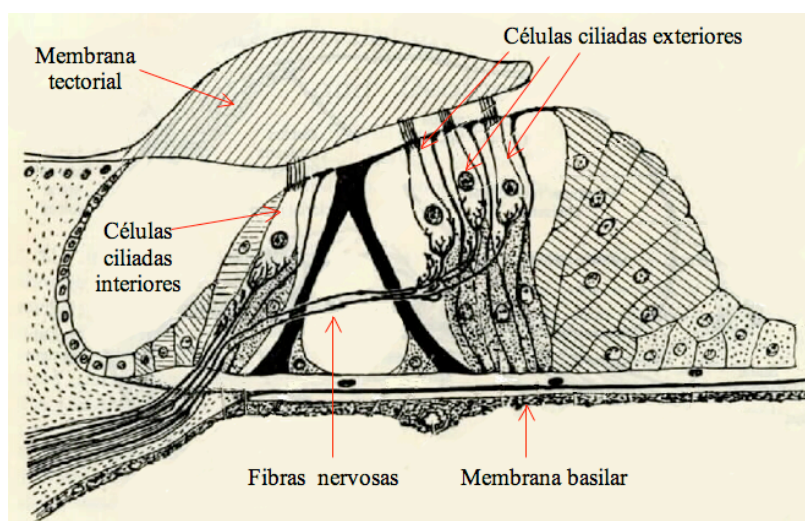


Figura 5. Representação do Órgão de Corti. Adap. de Maroonroge, Emanuel e Letowski (2009, p. 292).

O órgão de Corti situa-se no canal médio. É nele que se localizam as células ciliadas. Quando agitadas por vibrações mecânicas produzem os impulsos nervosos enviados para o cérebro. Existem dois tipos de células ciliadas: as células ciliadas exteriores e as células ciliadas interiores. Existem três filas com aproximadamente 12000 células ciliadas exteriores (Warren 2008, p. 10; Plack 2004, p. 8). Estas, possuem filamentos musculosos. Quando são estimulados, os filamentos contraem-se produzindo pequenos sons. Esta é a causa dos conhecidos zumbidos ou tinidos (*tinnitus*).

Segundo Mathews (1999, p. 9) as células ciliadas exteriores têm como função compensar possíveis perdas de energia mecânica, que possam acontecer ao longo da membrana basilar e por outro lado, controlar a intensidade sonora através da diminuição da vibração sobretudo, quando nos expomos a sons de elevada intensidade. Embora sejam em muito maior número do que as células ciliadas interiores recebem apenas 5% das enervação das fibras do nervo auditivo.

As células ciliadas interiores distribuem-se ao longo de uma fila com aproximadamente 3500 células (Warren 2008, p. 10; Plack 2004, p. 8). Estas células recebem 95% das enervações das fibras do nervo auditivo, sendo elas as principais responsáveis pela produção da sensação de audição. Quando danificadas provocam perdas auditivas acentuadas e irreversíveis.

As células ciliadas estão cobertas por uma membrana delicada, flexível e gelatinosa que se chama membrana tectorial. Os cílios (tufos parecidos com cabelo que se

estendem à superfície das células ciliadas) estão encaixados na membrana tectorial. Quando um estímulo sonoro provoca oscilações na membrana basilar, a membrana tectorial move-se estimulando os cílios.

Para compreendermos melhor o fenómeno da audição é importante saber como é que a membrana basilar responde às vibrações dos ossos vestibulares, que empurram a janela oval para dentro e para fora. Este aspecto mecânico, torna-se mais fácil de entender, se se desenrolar a cóclea de maneira a adoptar uma forma cónica (Figura 6).

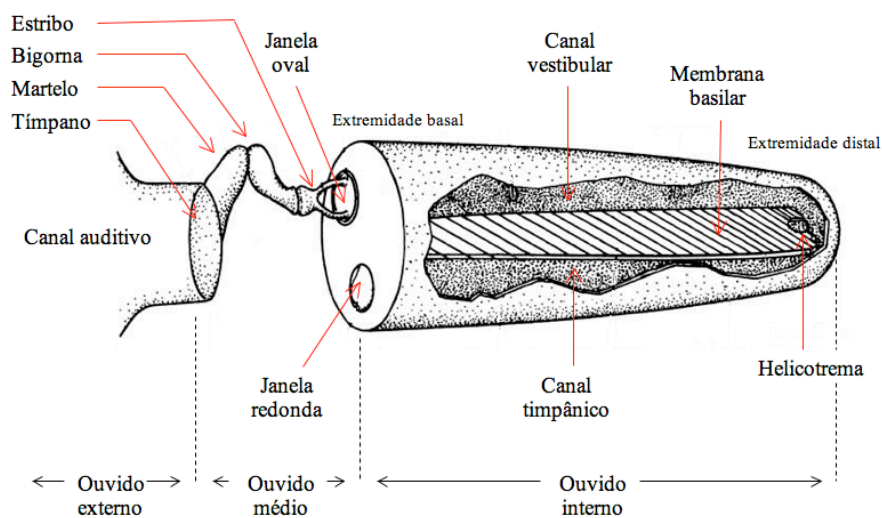


Figura 6. Representação cónica da cóclea. Adap. de Warren (2008, p. 8).

A largura da membrana basilar aumenta desde a extremidade basal até à extremidade distal. No entanto, apesar de ser mais estreita, a extremidade basal apresenta maior rigidez do que a extremidade distal. Segundo Hall (1990, p. 93) as vibrações provocadas pelas altas frequências estimulam sobretudo, a região junto à janela oval, por ser mais estreita e mais rígida. Pelo contrário, as vibrações originadas pelas baixas frequências estimulam a região junto da *helicotrema*, por ser mais ampla e mais flexível. Ao estimular a membrana basilar com um som puro (i.e., uma onda sinusoidal) a janela oval é empurrada e a membrana basilar produz uma onda vibratória que se propaga desde a janela oval em direcção à *helicotrema*. A membrana basilar funciona como um filtro de frequências (Plack 2004, p. 7). A Figura 7, mostra a distribuição das frequências sonoras na cóclea ao longo da membrana basilar.

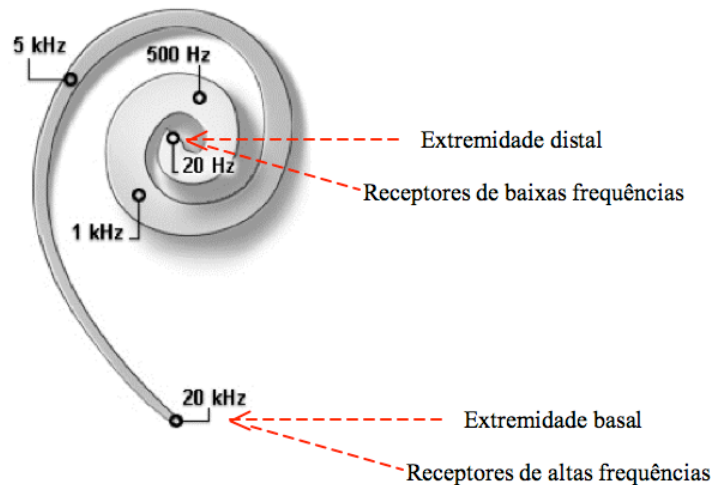


Figura 7. Distribuição das frequências sonoras ao longo da cóclea e da respectiva membrana basilar. As frequências sonoras distribuem-se de forma descendente (i.e., das altas para as baixas frequências), desde a extremidade basal até a extremidade distal da membrana basilar. Adap. de Warren (2008, p. 12).

Os receptores de altas frequências encontram-se na extremidade mais estreita (basal) junto da janela oval e os receptores de baixas frequências na extremidade mais larga (distal) (Plack 2004, p. 7).

A vibração produzida pelas ondas sonoras propaga-se até determinado limite. A onda produzida por um som puro não ultrapassa o seu limite de propagação. Este fenómeno de corte, que Mathews (1999, p. 7) designa de “*cutoff point*”, explica porque é que uma onda sonora de alta frequência não consegue “mascarar” ou ocultar uma onda sonora de baixa frequência, enquanto que uma onda de baixa frequência pode mascarar uma onda de alta frequência.

A título de resumo, realçamos as seguintes ideias (Mathews 1999, pp. 9-10):

- A membrana basilar realiza uma análise espectral aos sons.
- Diferentes frequências sonoras estimulam diferentes locais da membrana basilar. Dois estímulos auditivos que se encontram próximos em frequência, perturbam mais ou menos a mesma região, mas se se encontrarem significativamente afastados cada um deles estimula uma região diferente.
- Uma frequência sonora pode mascarar ou ocultar outra frequência que se encontre próxima, mas não consegue mascarar ou ocultar uma frequência que se encontre significativamente afastada dela.

- Sons de baixa frequência podem mascarar sons de alta frequência, mas sons de alta frequência não conseguem mascarar sons de baixa frequência (Pierce e Schubert 1987, p. 7).
- Sons puros separados por frequências muito afastadas são percebidos como sons distintos e diferenciados.
- Sons próximos em frequência, ao invés de serem percebidos como distintos, interagem resultando na percepção de apenas um som contendo batimentos, “rugosidade” ou “aspereza” sonoras, ou seja, a percepção de um efeito de modulação de amplitude (Scharine, Cave e Letowski 2009, p. 417; Warren 2008, p. 103; Bregman 1990, p. 236). O intervalo, no qual esta percepção de interferência ocorre denomina-se banda crítica (Warren 2008, p. 72; Mathews 1999, p. 9; Bregman 1990, p. 504).

I.2 Cérebro

É no cérebro que ocorrem as etapas finais do sistema auditivo. É nele que se encontram alojados os centros auditivos. Os impulsos nervosos originados a partir da estimulação das células ciliadas (ouvido interno) viajam até ao cérebro através do nervo auditivo, e deste, até ao córtex cerebral.

I.2.1 Nervo Auditivo

O nervo auditivo é constituído pelo nervo coclear, que transmite informação auditiva e pelo nervo vestibular, que transmite informação sobre o movimento e o equilíbrio da cabeça.

Os impulsos nervosos são transmitidos através do nervo coclear até aos centros auditivos do tronco encefálico e do córtex cerebral.

I.2.2 Córtex Cerebral

O córtex cerebral é um tecido fino composto essencialmente por uma rede de células nervosas (neurónios) densamente interligadas (Figura 8).

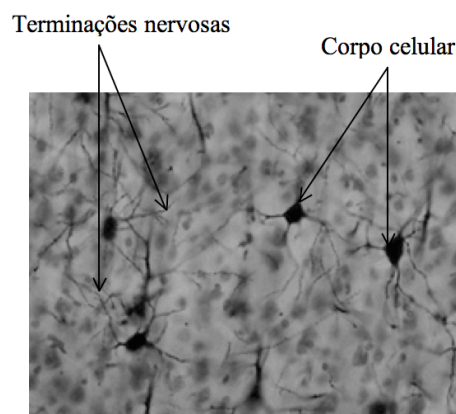


Figura 8. Células nervosas (neurónios) distribuídas no córtex cerebral.⁵

As terminações nervosas dos neurónios encontram-se à superfície do cérebro. Os axónios (o filamento que vai desde o corpo celular até aos ramos terminais do neurónio), espalhados por baixo da superfície interligam as várias regiões cerebrais.

Segundo Mathews (1999, p. 18) se se aumentar a frequência de um som puro, a excitação no córtex cerebral descreve um caminho padronizado, que se desloca para cima e para a direita do córtex cerebral.

I.2.3 Células Nervosas (Neurónios)

As células nervosas ou neurónios são células especializadas na condução de mensagens electroquímicas (i.e., impulsos nervosos) de uma parte para outra do corpo. Juntas, num feixe formam os nervos.

Em termos gerais, os neurónios são constituídos pelos seguintes elementos: corpo celular (que contém o núcleo da célula), dendrites (ramos proximais), nódulos de *Ranvier*, axónio e telodendrites (ramos terminais) (Figura 9).

⁵ Adap. de <http://brainmaps.org/index.php?p=screenshots>

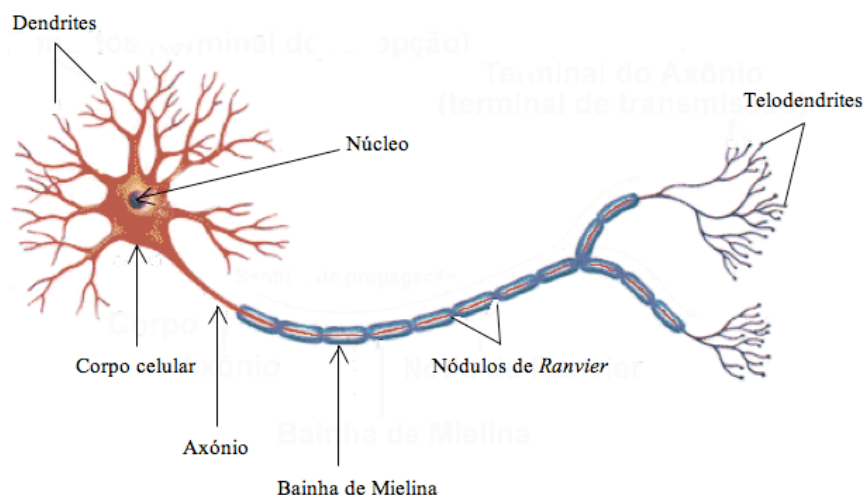


Figura 9. Constituição geral de um neurónio. Adap. de Mathews (1999, p. 12).

Os neurónios diferem no tamanho, no número de dendrites, de telodendrites, no comprimento dos axónios, e de acordo, com a função que desempenham (e.g., funções sensitivas, funções de associação ou funções motoras).

I.2.4 Transmissão dos Impulsos Nervosos

Quando um neurónio recebe um estímulo modifica as suas características produzindo uma pequena corrente eléctrica. Esta, circula num único sentido, das dendrites para o corpo celular e deste para o axónio. O espaço entre a dendrite de um neurónio e as telodendrites de um neurónio adjacente chama-se Fenda Sináptica. É nesta fenda, que se realizam as sinapses nervosas, ou seja, a transmissão dos impulsos nervosos entre os neurónios. Nas extremidades das telodendrites existem vesículas com substâncias químicas (os neurotransmissores), que são libertadas na Fenda Sináptica. Os neurotransmissores vão ligar-se a receptores existentes nas dendrites fazendo passar o impulso nervoso para um neurónio adjacente.

De acordo com Mathews (1999, p. 11) os tecidos que se encontram no exterior dos neurónios possuem uma carga eléctrica positiva em relação ao interior dos mesmos. Ou seja, em estado de repouso os neurónios encontram-se com uma carga eléctrica negativa, aproximadamente, 50 milivolts em relação aos tecidos exteriores circundantes.

Alguns tipos de axónios encontram-se revestidos por uma camada isolante de gordura, denominada Bainha de Mielina. A Bainha de Mielina tem como função

conservar a energia e a velocidade de propagação dos impulsos nervosos ao longo dos axónios.

As interrupções na Bainha de Mielina denominam-se Nódulos de *Ranvier*. Estas interrupções colocam os axónios em contacto com os tecidos circundantes polarizados com cargas eléctricas positivas. Quando um impulso nervoso alcança uma destas interrupções o isolamento nervoso quebra-se e é acrescentada energia ao impulso. Assim, o impulso nervoso, que viaja ao longo do axónio é constantemente regenerado mantendo a sua intensidade à medida que viaja ao longo do mesmo.

Depois de disparar o impulso nervoso o neurónio entra num estado refractário. Só dispara um novo impulso, após a sua carga ser regenerada, através de um processo de reajustamento químico, que acontece dentro do próprio corpo da célula nervosa.

Após o período refractário, o neurónio encontra-se pronto para disparar um novo impulso nervoso, quer espontaneamente, quer quando estimulado pelo impulso nervoso de outro neurónio adjacente, ou ainda, no caso das células ciliadas interiores com o movimento dos seus cílios.

I.2.5 Tempo de Propagação dos Impulsos Nervosos

De acordo com Mathews (1999, pp. 12) os impulsos nervosos viajam ao longo dos axónios a uma velocidade comparativamente mais lenta em relação à velocidade de propagação das ondas sonoras no ar. À temperatura ambiente de 20° graus Célsius e com a humidade relativa de 50%, as ondas sonoras viajam à uma velocidade de 344 m/s (Pierce 1999, p. 37).

A velocidade de transmissão dos impulsos nervosos aumenta linearmente de acordo com o diâmetro do axónio. A velocidade de propagação dos impulsos nervosos num axónio com um diâmetro aproximado de 2 micron (1 micron equivale a 1 milésimo de milímetro) é aproximadamente de 10 metros por segundo. Para um axónio com 20 micron de diâmetro a velocidade de propagação ronda os 120 metros por segundo. Estima-se, que a transmissão de uma mensagem eléctrica, desde a cabeça humana até aos pés demore sensivelmente um quinto de segundo (i.e., 200 milissegundos).

II PSICOACÚSTICA

Imaginem que nos encontramos no interior de um túnel, atrás e ligeiramente afastados de uma tela esticada feita com um lençol. Do outro lado da tela várias pessoas atiram bolas de ténis a diferentes distâncias. Podem atirar um número de bolas indeterminado e sempre que o entenderem fazer. A nossa tarefa consiste em tentar descobrir, apenas olhando para a maneira como a tela se move (para a frente e para trás), quantas pessoas existem, como são (i.e., se são altas, baixas, fortes), se estão perto ou longe, se estão paradas ou se caminham na nossa direcção, etc. Este é o principal “problema” que o nosso sistema auditivo enfrenta ao usar o tímpano como porta de entrada para a audição. A analogia descrita encontra-se representada na Figura 10.

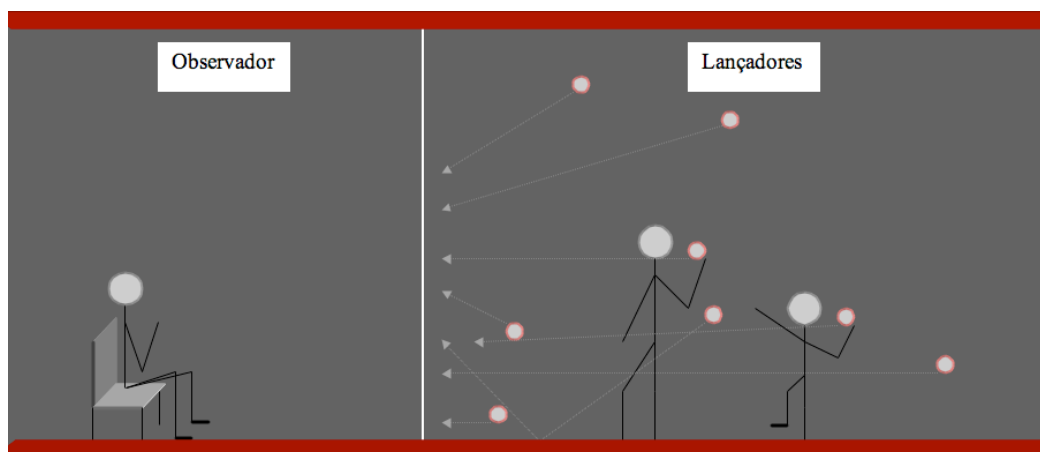


Figura 10. Analogia psicoacústica do tímpano.

O som é transmitido pelo ar através de moléculas que vibram com uma determinada frequência. Estas, ao embaterem contra o tímpano fazem-no vibrar com mais ou menos força para dentro e para fora (fases de compressão e de rarefacção). A força com que o fazem provoca alterações na percepção da intensidade sonora (*Loudness*). A velocidade com que o fazem provoca alterações na percepção da altura sonora (*Pitch*).

Não há nada nas moléculas de ar que nos informe de onde é que elas vieram nem com que objectos ou fontes sonoras estão associadas. Muitas vezes, a entrada do sinal é incompleta ou ambígua. Por exemplo, o som das vozes mistura-se com outras vozes, o som de máquinas com outros ruídos, o som do vento com o som do mar, o som dos nossos passos com o som dos passos de outras pessoas a caminhar, etc. Como é que o cérebro lida com isto? Como é que o cérebro lida com esta mistura de moléculas que embatem contra o tímpano?

No ouvido, as moléculas de ar atravessam o canal auditivo fazendo vibrar o tímpano. Este, pressiona os ossos vestibulares (martelo, bigorna e estribo). Os ossos perturbam o fluído existente na cóclea (é na cóclea que as vibrações são convertidas em impulsos electroquímicos). O fluído, por sua vez agita as células (cílios) (Pierce 1999, p. 89) e estas, emitem impulsos nervosos que viajam ao longo do nervo auditivo em direcção ao cérebro (Emanuel, Maroonroge e Letowski 2009, p. 307; Mathews 1999, p. 1).

Não há uma teoria que explique, por si só, como é que percebemos a altura dos sons. As três teorias principais que abordam esta temática são: a *Place Theory*, a *Temporal Theory* e a *Duplex Theory* (Cheveigné 2004). A *Place Theory* de Helmholtz (Licklider 1951, p. 128) preconiza que o ouvido interno funciona como um analisador de frequências e que o estímulo sonoro é decomposto em várias componentes sinusoidais, cada uma das quais estimula uma região diferente da membrana basilar. Por exemplo, as altas frequências estimulam a região basal, enquanto que as baixas frequências estimulam a região distal da respectiva membrana (Warren 2008, p. 16; Cheveigné 2004, p. 9). Para além disso, sugere que a percepção de altura sonora é determinada pela frequência fundamental do som (Cheveigné 2010, p. 84). No entanto, esta teoria tem dificuldade em explicar a percepção de altura sonora de sons cuja frequência fundamental se encontra omissa (*missing fundamental*). Ou seja, apesar da frequência fundamental não ter sido fisicamente gerada ou porque foi previamente retirada do som, conseguimos mesmo assim, percebê-la (Cheveigné 2004, p. 9).

A *Temporal Theory* de Fletcher (Licklider 1951, p. 128) tenta explicar a percepção da altura dos sons através da detecção de padrões periódicos realizados durante as descargas electroquímicas das fibras nervosas (Cheveigné 2004, p. 69). Ao contrário da teoria anterior, a *Temporal Theory* explica o fenómeno da *missing fundamental* da seguinte maneira: como os padrões periódicos não resultam da acção isolada de cada uma das componentes sinusoidais, mas sim da sua interacção, mesmo que a frequência fundamental seja omissa a sensação de altura sonora mantém-se (Warren 2008, p. 75). No

entanto, esta teoria, também revela algumas fragilidades sobretudo, na sua incapacidade de demonstrar, de forma inequívoca a extracção de um único impulso por período. Esta dificuldade torna-se ainda maior, quando se trata de sons complexos, por estes apresentarem mais do que um impulso por período (Cheveigné 2010, p. 85).

Por último, temos a *Duplex Theory* de Licklider (1951) que incorpora as duas teorias anteriores, na tentativa de criar um modelo mais abrangente e capaz de colmatar as fragilidades anteriormente mencionadas. Assim, a *Duplex Theory* preconiza que na determinação da altura sonora dos sons o sistema auditivo realiza por um lado, uma análise espectral (*Place Theory*) e por outro lado, uma análise de auto-correlação entre as frequências e os intervalos de tempo que medeiam os estímulos sonoros (*Temporal Theory*). Ou seja, o sistema auditivo realiza uma análise espectral às frequências, ao nível do ouvido interno e uma análise temporal ao comportamento dos impulsos nervosos, ao nível do sistema nervoso central (Licklider 1951, p. 128).

De acordo com Lehrer (2006)⁶ se o estímulo nervoso que chega ao cérebro for suave e regular, ouvimo-lo como um som geralmente agradável, por exemplo um intervalo de quinta perfeita. Pelo contrário, quando o som que desencadeia o estímulo nervoso é irregular, arritmico e imprevisível é percebido como algo de que não gostamos, por exemplo um intervalo de segunda menor.

Quisemos testar, se existe uma relação directa entre a percepção auditiva dos intervalos musicais mencionados e a sua representação gráfica em forma de onda sonora. Ou seja, constatar se o intervalo de quinta perfeita apresenta ou produz uma forma de onda graficamente mais regular do que o intervalo de segunda menor. Assim, começou-se por construir em síntese de som digital os dois intervalos mencionados.

Depois de obtidos os ficheiros áudio dos respectivos intervalos musicais, extraíram-se as representações gráficas das formas de onda. As formas de onda dos intervalos encontram-se graficamente representadas na Figura 11 e Figura 12.

⁶ Musical Language (Radio Lab: 21 April 2006) WNYC, New York Public Radio em <http://www.wnyc.org/shows/radiolab/episodes/2006/04/21#>

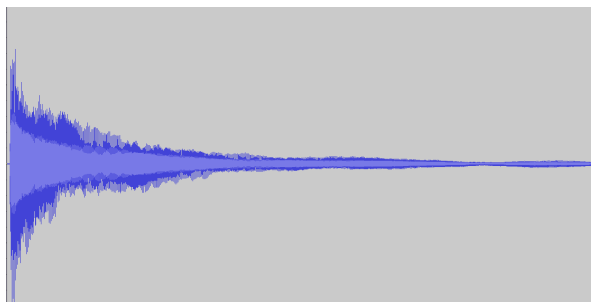


Figura 11. Representação da forma de onda de um intervalo de quinta perfeita (C3-G3) tocado em uníssono. O intervalo foi obtido através de síntese de som digital de dois sons puros de 130.8 Hz e 195.9 Hz.

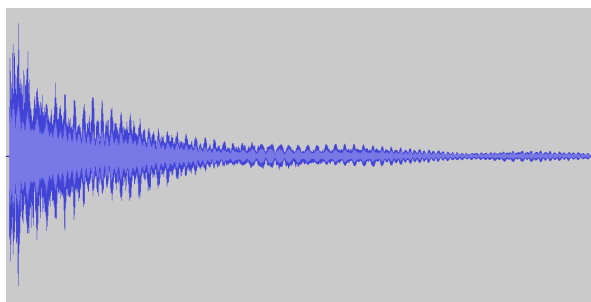


Figura 12. Representação da forma de onda de um intervalo de segunda menor (C3-Db3) tocado em uníssono. O intervalo foi obtido através de síntese de som digital de dois sons puros de 130.8 Hz e 138.5 Hz.

Observa-se em ambas as representações gráficas das formas de onda dos intervalos musicais seleccionados um decaimento de amplitude aproximadamente exponencial. No entanto, na forma de onda do intervalo de segunda menor (Figura 12), parece existir uma modulação de amplitude periódica, eventualmente resultante de um efeito de batimento entre as frequências.

Segundo Helmholtz (1895, p. 165) a percepção da dissonância ocorre quando acontecem batimentos entre os parciais. Estes batimentos surgem quando dois sons se encontram separados por um intervalo relativamente pequeno, como é caso do intervalo de segunda menor, por nós escolhido.

Hall (1990, p. 384) explica o fenómeno dos batimentos da seguinte maneira: suponhamos que estamos a ouvir simultaneamente duas ondas sinusoidais de igual amplitude. Se ambas possuírem a mesma frequência e se se encontrarem em fase, obtém-se uma onda sinusoidal com o dobro de amplitude. No fundo, ouvimos o mesmo som, ou seja, a mesma altura sonora, mas de maneira mais intensa. Se as duas frequências forem muito diferentes ouviremos dois sons puros distintos. Mas, se as frequências forem ligeiramente diferentes ouviremos um único som com batimentos, cuja frequência

resultante (f_0) é igual à média das frequências (Equação 1) (Scharine, Cave e Letowski 2009, p. 417).

$$f_0 = (f_1 + f_2)/2 \quad (1)$$

À medida que a separação entre as frequências aumenta, os batimentos tornam-se cada vez mais rápidos. Mesmo que não os consigamos captar individualmente continuaremos a perceber a sua presença através da rugosidade ou aspereza sonoras por eles produzida.

À medida que as duas frequências se aproximam dos limites de uma determinada banda crítica de separação, começamos a ouvir dois sons distintos. Quando as frequências se encontram suficientemente separadas uma da outra, de maneira a que possam estimular diferentes regiões da membrana basilar, a sensação de rugosidade ou de batimentos desaparece (Figura 13).

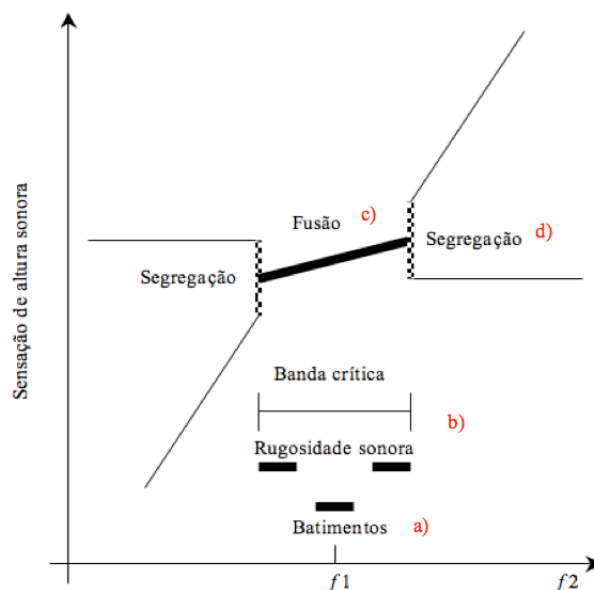


Figura 13. Combinação de duas ondas sinusoidais; f_1 é a frequência fixa e f_2 a frequência variável. Quando as duas frequências se encontram muito próximas originam batimentos (a). Dentro de uma determinada banda crítica percebemos uma sensação de rugosidade sonora (b). Dentro da área representada pela letra (c) ouvimos praticamente um único som (fusão sonora). Fora desta área, começamos a ouvir com clareza dois sons puros distintos (d). Adap. de Hall (1990, pp. 384).

A Frequência de batimentos (f_b) é dada pela Equação 1.1 (Scharine, Cave e Letowski 2009, p. 417; Warren 2008, p. 68; Loy 2006, p. 173; Cheveigné 2004, p. 26; Pierce e Schubert 1987, p. 11) .

$$f_b = f_2 - f_1 \quad (1.1)$$

Por exemplo, se $f_1 = 110$ Hz e $f_2 = 115$ Hz, a frequência de batimentos será igual a cinco Hertz, ou seja, cinco batimentos por segundo.

Mathews (1999, p. 8) reforça esta ideia ao afirmar que um som composto por duas frequências, que não se encontram muito separadas uma da outra, ao estimularem simultaneamente as mesmas células auditivas (i.e., os cílios das células ciliadas) originam batimentos ou rugosidade sonora. À medida que aumenta a distância entre as duas frequências começa por ouvir-se uma vibração. Se se aumentar ainda mais a distância entre elas, a partir de um determinado momento ouvimos dois sons sinusoidais claramente distintos.

No entanto, a dissonância não se deve somente à existência de batimentos entre as frequências. Se se aumentar a distância entre as frequências do intervalo de segunda menor (C3, 130.8 Hz e Db6, 1108.7 Hz) os batimentos desaparecem, mas a dissonância permanece. Isto, deve-se ao facto, das frequências dos parciais em causa não fazerem parte da mesma série harmónica, ou seja, não existe uma relação múltipla ou submúltipla inteira entre elas.

Esta perspectiva de percepção de intervalos musicais com altura definida, refere-se exclusivamente a intervalos musicais constituídos por sons puros tocados em uníssono.

Nem todos os intervalos de segunda são percebidos de forma desagradável ou dissonante, tudo depende do material sonoro que os antecede e do que se vai passar musicalmente a seguir. No entanto, parece haver uma relação entre o tipo de impulso nervoso que o som produz no cérebro e como sentimos esse som. De acordo com Lehrer a fronteira entre as pessoas que sentem o som de forma agradável ou desagradável é muito maleável.

“O que realmente ouvimos é o fim de uma longa cadeia de eventos mentais que dão origem a uma impressão – imagem mental – do mundo físico que nos rodeia. Em nenhuma outra dimensão isto se torna tão evidente, como na situação em que o nosso cérebro tenta impor uma ordem e uma estrutura a uma sequência de sons para criar aquilo a que chamamos Música. ” - Levitin (2008, p. 34).

III PSICOLOGIA COGNITIVA

Para além da Psicoacústica e das Neurociências, a Psicologia Cognitiva é uma área de investigação e de conhecimento, que se tem interessado pela percepção dos sons e da música. Os psicólogos da cognição estão interessados sobretudo, na representação final e interna dos eventos sonoros.

Existe um mundo inteiro para ser entendido e todos os sentidos fornecem informação ao observador sobre esse mundo. No entanto, o enfoque do nosso estudo concentrar-se-á sobretudo, sobre a audição. Quando necessário, socorrer-nos-emos de analogias, ou seja, de processos e princípios visuais para descrever ou explicar alguns fenómenos auditivos.

Por vezes, os processos podem ser partilhados por ambos os sentidos, mesmo quando os estímulos não são os esperados. De acordo com Gregory (1968, p. 280), isto acontece quando um determinado órgão sensorial assinala “um tipo de estímulo errado”. Por exemplo, ao pressionarmos os ouvidos simultaneamente com ambas as mãos enviamos sinais (i.e., corrente eléctrica) para o cérebro, onde serão percebidos como som. O mesmo acontece, quando aplicamos este procedimento aos olhos. Se pressionarmos simultaneamente os olhos com ambas as mãos enviamos sinais (i.e., corrente eléctrica) para o cérebro, onde serão interpretados como luz.

Temos ainda os exemplos dos fenómenos de Constância de tamanho (*Size constancy*) na visão e de Constância de percepção de intensidade sonora (*Loudness constancy*) na audição. O fenómeno de Constância de tamanho é revelado através do facto de conseguirmos perceber a dimensão real dos objectos independentemente da distância a que se encontram de nós (Shepard 1999, p. 25). O fenómeno de Constância de percepção de intensidade sonora pode ser descrito da seguinte maneira: quando um músico diminui a intensidade sonora produzida pelo seu instrumento, aparentemente deveria transmitir ao ouvinte a sensação que o som vem de longe. No entanto, o ouvinte possui a capacidade de perceber que o som produzido nessa circunstância não vem de

longe, mas do facto, de o músico o estar a tocar com menor intensidade (Shepard 1999, p. 25). Isto, acontece, porque para além da intensidade, o ouvinte tem também em consideração a distribuição do conteúdo espectral do som. Por exemplo, um som musical tocado forte, mas longínquo, será mais rico em harmónicos superiores do que o mesmo som, tocado piano e próximo. As altas frequências das componentes do som tendem a aumentar ou a diminuir, de acordo com o vigor com que o músico toca o seu instrumento (Shepard 1999, p. 26). Para além disso, como as baixas frequências sofrem menor atenuação sonora do que as altas frequências, um som longínquo será também mais rico em harmónicos inferiores (Scharine, Cave e Letowski 2009, p. 465; Loy 2006, p. 194). Outra pista importante na determinação da distância entre a fonte sonora e o ouvinte (em espaços fechados ou em situações que produzam eco) encontra-se na relação entre o som directo e o som reflectido (Reverberação) (Warren 2008, p. 59; Pierce 1999, p. 97). Se a fonte sonora se encontrar perto do ouvinte, o som directo é mais intenso do que o som reflectido e consequentemente o som emitido tornar-se-á mais nítido, e vice-versa (Scharine, Cave e Letowski 2009, p. 465; Shepard 1999, p. 26).

Embora não seja central neste trabalho, mas uma vez que mencionámos as analogias entre a audição e a visão, não podemos de deixar de ressaltar que ambos os sentidos possuem funções e comportamentos diferentes, podendo em certas situações interagir. Vejamos o exemplo do Efeito McGurk (McGurk e MacDonald, 1976), ou o exemplo da indução de uma ilusão visual a partir de som, designada por *Sound-Induced Illusory Flash* (Kamitani et al., 2002). Em termos gerais, o Efeito McGurk consiste no seguinte: num filme previamente preparado, a sílaba [ga], gesticulada através dos lábios e dobrada pelo som [ba] é ouvida pelos observadores, como [da]. Com esta ilusão auditivo-visual, os autores demonstraram a influência da visão sobre a percepção da fala (McGurk e MacDonald 1976, p. 746). A *Sound-Induced Illusory Flash* baseia-se no seguinte: quando um único *flash* de luz é acompanhado por múltiplos *beeps* sonoros este é percebido como múltiplos *flashes*. Com esta ilusão, os autores demonstraram que o som pode alterar a percepção visual (Kamitani et al., p. 152).

IV PRINCIPAIS TEORIAS DA PERCEPÇÃO

O que é a percepção? Uma questão particular a considerar, é o processo pelo qual a energia física recebida pelos órgãos dos sentidos constitui a base da experiência perceptiva. A informação sensorial é convertida em representações visuais, auditivas, olfactivas, tácteis e gustativas. Ou seja, é convertida na percepção de mesas, computadores, flores, edifícios, automóveis, aviões, sons, cheiros, gostos e experiências tácteis.

Outra questão importante, consiste em saber em que medida a percepção depende directamente das informações presentes no estímulo. Alguns psicólogos da cognição, como Gibson (1966) argumentam que os processos perceptivos são directos, ou seja, não há processamento de informação, o que observamos é o que recebemos - teoria da percepção directa ou processamento ascendente (*bottom-up*). Outros, como Gregory (1970) pelo contrário, defendem que os processos são indirectos, ou seja, dependem das informações disponíveis no próprio estímulo e de fenómenos internos, como as expectativas, a motivação, a emoção e o grau de conhecimento prévio do observador - teoria construtivista, indirecta ou processamento descendente (*top-down*).

IV.1.1 Gibson e a Percepção Directa ou Processamento Ascendente

Gibson opôs-se à ideia de que a percepção envolve um processamento descendente. Criticou sobretudo, a discussão à volta das ilusões visuais. Argumentou, que as ilusões visuais são exemplos artificiais e não imagens encontradas num ambiente visual natural. No entanto, esta convicção de Gibson, pode ser refutada recorrendo à existência de ilusões com origem em processos naturais (i.e., ilusões físicas). O exemplo de uma ilusão, que ocorre devido às propriedades físicas do meio ambiente e dos seus efeitos sobre o comportamento da luz (essencialmente ocorridos antes da luz atingir a retina) encontra-se no fenómeno das miragens. As miragens são ilusões ópticas muitas vezes

associadas a um Oásis no meio do deserto. São causadas pela reflexão da luz fora de uma fina camada de ar quente junto ao solo aquecido. Nos desertos, o reflexo do céu, fora dessa camada de ar poderá transmitir à distância a ilusão de um lago.

Gibson argumentou que a percepção é directa. Para ele, a sensação é a percepção. O que vemos é o que recebemos. Não há necessidade de transformação. As informações que recebemos sobre o tamanho, a forma e a distância são suficientemente detalhadas para que possamos interagir directamente com o meio ambiente.

No entanto, para nós o conhecimento é efectivamente imprescindível na percepção, porque muitas das propriedades indirectas dos objectos importantes na nossa interacção com o meio ambiente não são percebidas directamente através dos olhos, como a dureza e o peso, se está quente ou frio, se é comestível ou venenoso, etc. Por isso, consideramos ter sido um passo importante afirmar que as percepções são hipóteses. Desta maneira, podemos prever as características indirectas dos objectos e anteciparmos as nossas decisões de acção.

Um dos contributos atribuído a Gibson é o conceito de pistas visuais (*visual affordances*) (Goldstein 1981, p. 193). As pistas visuais são sugestões ou possibilidades de acção, que o meio ambiente nos fornece e que nos auxiliam na percepção. Por exemplo, na paralaxe de movimento, movimentos por parte do observador modificam as imagens em cada retina e as diferenças entre os sucessivos pontos de vista são usadas para inferir as distâncias. A paralaxe de movimento reflecte-se na audição da seguinte maneira: se a fonte sonora se encontra longe, os movimentos da cabeça do ouvinte interferem muito pouco na localização lateral dos sons. No entanto, se a fonte sonora estiver próxima, o movimento da cabeça provocará alterações na localização lateral dos sons e na percepção da intensidade sonora, para além, de simultaneamente, criar uma sensação imaginária de que a fonte sonora se está a movimentar (Scharine, Cave e Letowski 2009, p. 465). A paralaxe de movimento na audição ajuda-nos a compreender o efeito *Doppler*. Este efeito, provoca alterações na percepção auditiva da altura e da intensidade sonoras originadas, quer por fontes sonoras em movimento (e.g., quando um carro passa por nós), quer por movimentos do ouvinte em relação a uma fonte sonora fixa (e.g., passar junto a uma queda de água). De acordo, com a primeira situação, à medida que a fonte sonora se aproxima do ouvinte, as distâncias entre as fases de compressão e rarefacção das ondas sonoras são comprimidas, fazendo com que as frequências aumentem em altura em relação às frequências do som emitido, acompanhadas por um aumento progressivo da intensidade sonora. Depois da fonte sonora passar pelo ouvinte, as distâncias entre as

fases de compressão e rarefacção são expandidas, fazendo com que as frequências, que chegam ao ouvinte, diminuam em altura em relação às frequências do som emitido, para além de se verificar uma diminuição progressiva da intensidade sonora (Scharine, Cave e Letowski 2009, p. 431). Por outro lado, movimentos do ouvinte (aproximar-se e afastar-se) da fonte fonora ou do evento sonoro despoletado, também modificam a percepção auditiva em termos tímbricos e em termos de intensidade sonora. Podemos experimentar este fenómeno colocando água a correr numa torneira. Se aproximarmos e afastarmos a cabeça repetidamente da água a correr, percebemos as alterações que ocorrem, quer ao nível da intensidade, quer ao nível do timbre (i.e., o som da água torna-se mais intenso e mais agudo quando nos aproximamos).

Também a sobreposição representa um pista visual. Se a imagem de um objecto bloqueia ou se sobrepõe à imagem de outro objecto, o primeiro é visto como estando mais próximo (Figura 14).

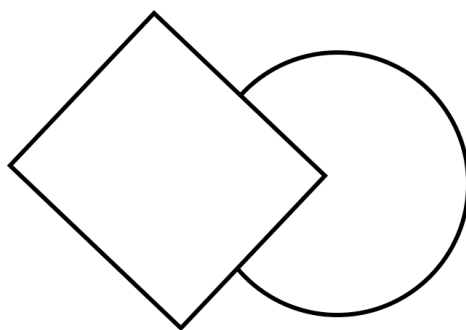


Figura 14. Sobreposição. Adap. Bregman (1990, p. 628).

O mesmo acontece no plano auditivo. Se dois sons são apresentados simultaneamente, mas com grandes diferenças de intensidade, o ouvinte ouvirá o som mais intenso como sendo aquele que se encontra mais próximo.

Ao nível da audição, existem pistas que nos permitem orientar no espaço acústico envolvente. Para além das pistas binaurais (i.e., quando os sons são escutados por ambos os ouvidos) existem também pistas auditivas monoaurais (i.e., quando os sons são escutados só por um ouvido). As pistas binaurais (e.g., a diferença de intensidade e a diferença temporal interaurais mencionadas no Sub-subcapítulo I.1.1, desta dissertação) fornecem-nos informações sobre a localização dos sons no plano horizontal e ao longo do eixo x (lateralização ou azimuth). Ou seja, ajudam-nos a perceber se o som vem do lado

direito ou do lado esquerdo em relação ao plano mediano do ouvinte (Scharine, Cave e Letowski 2009, p. 391; Warren 2008, p. 35). Por sua vez, as pistas monoaurais dão-nos informações sobre a localização dos sons no plano vertical e ao longo dos eixos *y* e *z* (elevação ou zénite). Ou seja, permitem-nos distinguir se os sons vêm pela frente, por trás, por baixo ou por cima do ouvinte (Scharine, Cave e Letowski 2009, p. 460; Warren 2008, p. 35).

Apesar de Gibson ter alegado, que as ilusões usadas em trabalhos experimentais são situações artificiais construídas em laboratório e que tais situações perceptivas são improváveis de ser encontradas no mundo real, a sua teoria directa ou processamento ascendente não explica porque é que às vezes as percepções são imprecisas ou distorcidas, como acontece nas ilusões.

Gibson ao afirmar, que a percepção é um processo directo, onde não é necessário realizar inferências e que não depende dos nossos conhecimentos previamente adquiridos, nega categoricamente a evidência do fenómeno das ilusões cognitivas.

IV.1.2 Gregory e a Percepção Construtivista, Indirecta ou Processamento

Descendente

Gregory afirma que a percepção é um processo construtivo. Depende, sobretudo, do processamento cerebral dos estímulos sensoriais. Percepcionar implica fazer inferências para formular a melhor hipótese sobre o que estamos a ver, ouvir ou sentir. Segundo ele, o conhecimento prévio e a experiência do passado são fundamentais na percepção. Quando olhamos para algo, desenvolvemos uma hipótese perceptiva que se baseia no conhecimento prévio adquirido. Estamos, activa e dinamicamente a construir a nossa percepção da realidade. Os receptores sensoriais recebem informações do ambiente, que depois são combinadas com as informações anteriormente armazenadas sobre o mundo que construímos ao longo da nossa existência. A formulação de hipóteses incorrectas pode levar a erros de percepção, como no caso das ilusões. Vejamos o exemplo da Ilusão Visual do Cubo de Necker (1832) (Figura 15).

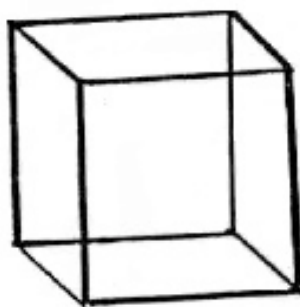


Figura 15. Ilusão Visual do Cubo de Necker. Adap. de Gregory (1977, p. 192).

Quando olhamos para o cubo, a sua orientação pode mudar repentinamente. O cubo torna-se instável. Um único padrão físico pode produzir duas percepções. Gregory argumentou, que este objecto parece mudar entre duas orientações, porque no cérebro se desenvolvem duas hipóteses igualmente plausíveis, sendo este incapaz de se decidir por uma delas. Quando a percepção muda não há alteração na entrada sensorial o que prova que a alteração na percepção ou na orientação do cubo não se deve à entrada do sinal sensorial, mas sim ao processamento descendente realizado. A Figura 16, ilustra através da explicitação da face anterior do cubo (face opaca), a mudança que acontece na orientação do mesmo, quando olhamos fixamente para ele durante algum tempo.

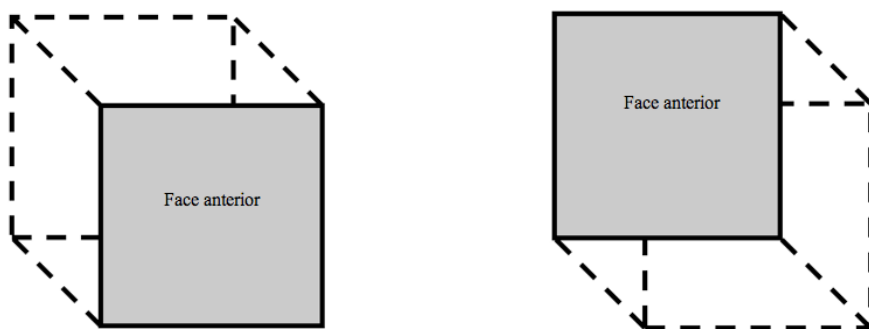


Figura 16. Posição inicial do cubo (imagem da esquerda). Posição posterior do cubo (imagem da direita).

É importante realçar, que esta ilusão visual também acontece com os cubos representados na Figura 16. Ou seja, se olharmos fixamente para qualquer um deles, a orientação inicial também se altera.

Warren e Gregory (1958, p. 612) criaram uma ilusão auditiva análoga à Ilusão Visual do Cubo de Necker, conhecida por *Loop Words*. Esta ilusão auditiva consiste no

seguinte: a palavra inglesa “say” repetida vezes sem conta e ininterruptamente muda para “ace”, depois regressa a “say”, e assim sucessivamente.

Se as percepções fazem uso de hipóteses para testar as informações sensoriais, podemos perguntar que tipo de hipóteses são e se podem ser modificadas pelo observador.

Segundo Gregory as hipóteses formuladas são como as hipóteses preditivas da ciência. Como os cientistas modificam as hipóteses de acordo com os fundamentos que encontram, parece que em alguns casos os observadores também o fazem. Por exemplo, se olharmos para a Figura 17 (imagem ambígua), provavelmente veremos na imagem o perfil de um rosto humano com óculos.



Figura 17. Imagem ambígua. Adap. de Goolkasian e Woodberry (2010, p. 170).

No entanto, há uma palavra escondida nele - *Liar*. A palavra está escrita na diagonal, desde o canto superior esquerdo até ao canto inferior direito.

A partir do momento que a palavra se torna visível ocorrem rápidas aprendizagens na percepção. Agora, a imagem ambígua contém sempre a palavra *Liar* quando olhamos para ela. Aprendemos a perceber o estímulo de uma maneira diferente.

Embora em alguns casos, como na Figura 17, exista uma relação directa entre a modificação das hipóteses e a percepção, noutros casos isso não é tão evidente. Por exemplo, porque é que as ilusões persistem mesmo quando temos conhecimento delas?

Seria de esperar, que o conhecimento que vamos adquirindo, pudesse de alguma maneira, alterar de forma adaptativa as nossas hipóteses. Esta relação entre a percepção e a aprendizagem é uma das lacunas que a teoria de Gregory não consegue explicar.

A abordagem construtivista sublinha o papel do conhecimento na percepção, no entanto, existem algumas perguntas por responder. Por exemplo:

- Se todos temos de construir o nosso próprio mundo com base em experiências, porque é que as percepções são tão semelhantes, mesmo em culturas diferentes?
- Porque é que após o nascimento, os recém-nascidos já demonstram a capacidade de reconhecer a inalterabilidade da forma? (Slater 1985, p. 62).
- Porque é que os recém-nascidos preferem a voz da mãe a outras vozes? (DeCasper e Fifer 1980, p. 1175).

A principal crítica feita aos construtivistas baseia-se, no facto, de terem subestimado a riqueza de estímulos sensoriais disponíveis no mundo físico real que nos rodeia em oposição ao trabalho de laboratório, onde grande parte das suas evidências têm sido alcançadas.

Construtivistas, como Gregory usam frequentemente o exemplo da inalterabilidade de tamanho para apoiar as suas explicações. Isto é, apercebemo-nos do tamanho de um objecto mesmo que a sua imagem na retina diminua à medida que o objecto se afasta (Constância de tamanho). No entanto, no mundo real as imagens da retina raramente são vistas de forma isolada, ao contrário do que acontece nas experiências de laboratório. Há uma grande variedade de informações sensoriais, incluindo outros objectos distantes e em movimento no horizonte. Esta fonte rica de informação sensorial transmitida pelo meio ambiente que nos rodeia é aproveitada por Gibson para explicar a sua teoria da percepção directa ou processamento ascendente.

Achamos que as duas teorias da percepção são conciliáveis e complementares. É verdade, que a informação ou os estímulos existentes são suficientes para percepcionarmos o mundo que nos rodeia, mas também é verdade, que por vezes, o nosso cérebro tem de deduzir ou restaurar informação omissa. Por exemplo, se num som complexo (i.e., um som constituído por mais do que uma componente) (Fugiel 2011, p. 415; Cheveigné 2010, p. 71; Deutsch 2010, p. 11; Scharine, Cave e Letowski 2009, p. 418; Nakajima *et al.* 1988, p. 7; Ueda e Ohgushi 1987, p. 1195; Burns 1981, p. 467-468; Shepard 1964, p. 2346) omitirmos a frequência fundamental (i.e., a primeira componente) continuamos a senti-la ou ouvi-la, apesar de ela não ter sido gerada. Este fenómeno, inicialmente denominado por *missing fundamental* foi apresentado por Harvey Fletcher

em 1924. Mais tarde, em 1962, Schouten substitui o termo por *Virtual Pitch*, ao referir-se a sons, cuja frequência fundamental se encontra omissa. No entanto, de acordo com Pierce (1999, p. 70) este fenómeno não tem relevância em termos de percepção auditiva. Ao escutarmos música, ninguém consegue discriminar se os sons em relação à percepção de altura sonora possuem ou não frequência fundamental. Isto, deve-se ao facto, da percepção de altura sonora não depender da presença física da frequência fundamental, mas da existência de uma região espectral predominante, da qual fazem parte o terceiro, o quarto e o quinto harmónicos (Warren 2008, p. 79). Para se ter a percepção de altura sonora não é necessário que a frequência fundamental esteja presente (Cheveigné 2010, p. 78; Pierce 1999, p. 59).

O fenómeno inverso também ocorre. Ou seja, num som puro (um som constituído por uma única componente, por exemplo um som sinusoidal) (Cheveigné 2010, p. 71; Warren 2008, p. 2; Bregman 1990, p. 83) é possível ouvir os harmónicos subsequentes (h_2 , h_3 , h_4 , etc.), apesar de não terem sido gerados. Assim, através deste processo de inferência, melhoramos os nossos níveis de percepção e de aprendizagem. Hipoteticamente, permitem-nos aumentar as capacidades de adaptabilidade a qualquer tipo de situação ou ambiente, quer seja rico ou pobre em estímulos.

Para Levitin (2008 p. 34) a cadeia de eventos mentais começa com um processo chamado extracção de informação. Por exemplo, a partir da música o cérebro extrai informação considerada básica ou de baixo nível usando redes neuronais especializadas que descodificam o sinal em termos de percepção de altura sonora, localização espacial, intensidade, duração, timbre, etc. Este processo de extracção de elementos básicos designado por processamento ascendente (teoria directa) ocorre, segundo este autor, nas regiões mais periféricas e filogeneticamente mais antigas do cérebro. Posteriormente, segue-se um processo chamado integração. As partes do cérebro superior (principalmente no córtex frontal) recebem a informação básica das regiões mais baixas do cérebro integrando-as na percepção global dos eventos. Este processo é designado de processamento descendente (teoria construtivista).

Como o nosso trabalho de investigação se vai centrar exclusivamente nas ilusões sonoras realizadas através de processos artificiais e em contexto de laboratório, sentimos uma maior afinidade com a teoria construtivista. Por outro lado, como a teoria de Gibson nega as ilusões cognitivas será nesta dissertação preterida em favor da teoria de Gregory.

Os construtivistas ao recorrerem às ilusões para estudar os mecanismos da percepção humana foram consolidando ao longo do tempo alguns dos princípios que hoje

em dia suportam o conhecimento que temos sobre os processos de construção e funcionamento das ilusões cognitivas. Ao trabalharem as ilusões sobretudo, as de origem artificial foram criando um espaço potencialmente fértil para a criatividade, em geral e para a música, em particular. Referimo-nos concretamente à possibilidade de aplicar os mecanismos das ilusões cognitivas sonoras em composição musical.

IV.2 Principais Mecanismos da Percepção

A interpretação que fazemos de uma imagem pode ser bastante diferente do estímulo físico recebido pela retina (receptor sensorial da visão) como se pode comprovar através da ilusão visual representada na Figura 18.

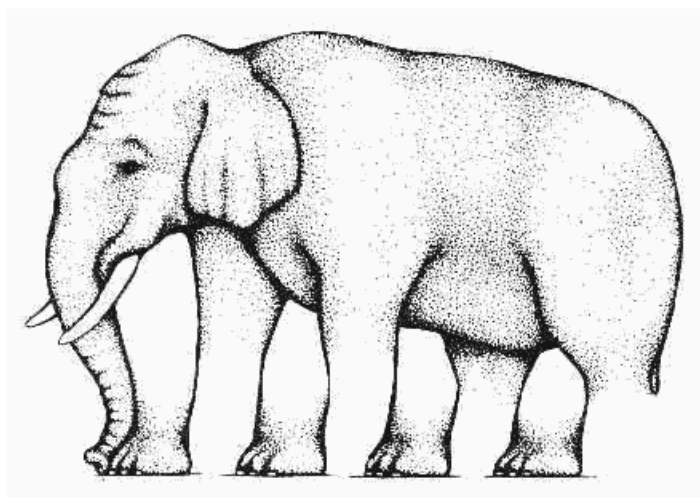


Figura 18. Ilusão Visual das pernas do elefante de Shepard. Adap. de Shepard (1990, p. 79).

A Ilusão Visual das pernas do elefante de Shepard (*Shepard's elephant legs*) consiste na representação bidimensional de um elefante, constituída por padrões de linhas e pontos. Os seres humanos têm tendência a interpretar esses padrões como objectos tridimensionais e constatarem automaticamente que o elefante possui mais patas do que o normal. Se fôssemos capazes de voluntariamente desligar ou apagar a representação cognitiva que fazemos da imagem verificávamos que o elefante possui somente quatro patas. Constatamos isso facilmente ao dirigirmos o olhar para a extremidade distal das patas ou para as unhas do elefante. O facto, de ser difícil de ver à primeira vez, somente as quatro patas prova que a nossa representação da imagem é bastante diferente do padrão

apresentado na retina. Segundo Shepard (1999, p. 23) o processo de interpretação e representação cerebral parece ser inconsciente, instintivo, automático e independente da nossa vontade.

Os mecanismos da percepção considerados mais importantes para o desenvolvimento do nosso estudo são: a inferência inconsciente, a percepção de preenchimento e os princípios gestálticos de agrupamento. De seguida, vamos abordar cada um dos mecanismos da percepção mencionados.

IV.2.1 Inferência Inconsciente

Helmholtz (1821-1894) foi um conceituado cientista alemão com contributos importantes para as Ciências Modernas, por exemplo na área da Acústica, através da realização de estudos sobre a percepção da altura do som.

Na Psicologia Cognitiva, Helmholtz ficou conhecido pela formulação do seu princípio de inferência inconsciente (*unconscious inference*). Este princípio permite-nos através da dedução, restaurar ou completar a informação incompleta ou insuficiente que nos chega aos sentidos (Cheveigné 2004, p. 3; Deutsch 1999, p. 299; Gregory 1997, p. 1; Bregman 1990, p. 225). A Figura 19 ilustra o princípio referido.

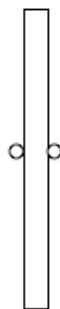


Figura 19. Inferência inconsciente. Somos capazes de reconhecer uma porta, apesar de só vermos um rectângulo longo e estreito quando a mesma se encontra entreaberta.

Para Levitin (2008, p. 35) quando o estímulo sensorial é a música, as inferências incluem vários factores para além dos próprios sons: o que veio antes da obra musical que estamos a ouvir, como começou, o que pensamos que virá a seguir, se o género ou estilo

da música nos é familiar, bem como qualquer informação adicional que tenhamos tido acesso previamente, através da leitura das notas do programa do concerto, etc.

O cérebro constrói uma representação sonora da realidade baseada nas características que compõem o que nós realmente ouvimos e as nossas expectativas do que achamos que deve ser ouvido.

Segundo Levitin (2008, p. 34) o cérebro enfrenta três dificuldades ao realizar os processos de extracção e integração da informação sonora, devido ao facto de esta poder ser indiferenciada, ambígua ou incompleta. Assim:

- a informação que chega aos receptores sensoriais pode ser indiferenciada, em termos de localização, origem e identidade;
- a informação será ambígua nos casos em que sons diferentes podem dar origem a padrões semelhantes de activação dos tímpanos, como no exemplo de dois sons com a mesma altura e intensidade tocados por dois instrumentos musicais distintos;
- a informação que recebemos raramente é completa uma vez que partes do som podem ser perdidas ou mascaradas por outros sons. O cérebro tem de deduzir o que realmente está a acontecer em termos sonoros.

A percepção auditiva é um processo de inferência. Inferência, não significa mais do que fazer uma interpretação provável a partir de dados incompletos. Assim, um sistema perceptivo que permite restaurar a falta de informação pode ajudar-nos a tomar decisões rápidas em situações de perigo. Mas também pode apresentar “fragilidades”. Por exemplo, o processo cognitivo de processamento descendente pode levar-nos a fazer percepções distorcidas da realidade através da redefinição da informação realizada ao nível do processamento ascendente. Para Levitin (2008, p. 35) isto é, em parte a base neuronal das ilusões cognitivas.

IV.2.2 Percepção de Preenchimento

Este mecanismo permite gerar informação de foma indirecta. Kanizsa (1913-1993) forneceu uma variedade de exemplos visuais sobre a percepção de preenchimento. Vejamos o exemplo representado na Figura 20.

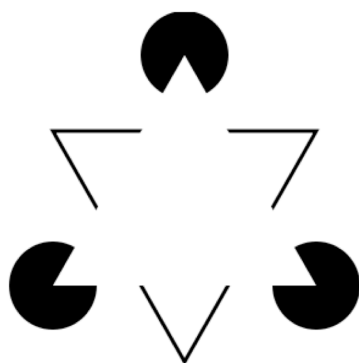


Figura 20. Triângulo de Kanizsa. Adap. de Kanizsa (1955, p. 16).

É possível constatar a presença de um triângulo branco localizado no centro da Figura 20. Contudo, o triângulo não foi previamente desenhado.

Em relação à percepção auditiva, a ocorrência deste fenômeno foi demonstrado, quer por Bregman (1996), quer por Warren (2008).

Bregman demonstrou este mecanismo da percepção através de *glissandos* ascendentes e descendentes de sons sinusoidais (1996, exemplo 29 – *perceptual continuation of a gliding tone through a noise burst*). Os sons foram quebrados por espaços vazios causando a percepção de interrupção. Quando os espaços vazios foram preenchidos por “explosões” controladas de ruído branco, os ouvintes inferiram que os *glissandos* eram contínuos.

Pode fazer-se a mesma experiência com música. Por exemplo, criar espaços vazios (silêncios) ao longo de um ficheiro áudio de uma música escolhida ao acaso, que se encontre armazenada no disco rígido do nosso computador. Ao escutar o resultado, os silêncios transmitem-nos a sensação de terem sido causados por uma interferência qualquer no sinal áudio. Se preenchermos os silêncios com ruído branco de elevada intensidade, em certas circunstâncias temos a percepção auditiva que a música continua a fluir de forma contínua.

Warren (2008, p. 162) demonstrou este mecanismo através de um exemplo chamado “*Temporal Induction of Speech: phonemic restoration*”. O exemplo consiste no seguinte: o fonema de uma palavra proferida num discurso é substituído por um ruído de elevada intensidade (e.g., uma tossidela). Geralmente, consegue-se ouvir a frase intacta, mesmo sabendo antecipadamente, que alguma da informação contida na frase vai faltar.

Este mecanismo de preenchimento não é apenas uma curiosidade de laboratório. Os compositores também exploram este mecanismo, pois sabem que a percepção de uma linha melódica vai continuar mesmo que parte dela seja obscurecida por outros

instrumentos. O mesmo acontece quando ouvimos as notas mais graves do piano ou do contrabaixo. Não estamos realmente a ouvir os 27.5 Hz ou os 35 Hz, porque estes instrumentos são normalmente incapazes de produzir muita energia nessas baixas frequências. Em vez disso, os nossos ouvidos preenchem ou restauram as informações dando-nos a ilusão de que a altura do som é efectivamente assim tão baixa (inferência inconsciente) (Levitin 2008, p. 35; Cheveigné 2004, p. 3).

IV.2.3 Princípios Gestálticos de Agrupamento

Os estudos da percepção realizados durante as décadas de 30 e 40 do século XX, deram origem à Escola da Psicologia Gestáltica. Esta escola centra-se na abordagem holística. Holismo é uma palavra grega que significa todo, inteiro ou total. É a ideia de que todas as propriedades de um determinado sistema, quer seja físico, químico, biológico, social, económico, psíquico, linguístico, musical, etc., não podem ser explicadas através dos seus elementos isolados. Em vez disso, o sistema como um todo determina como as partes se comportam. O todo é diferente da soma das suas partes (Palmer 2002, p. 179).

De acordo, com os fundadores da Psicologia Gestáltica, os psicólogos alemães Max Wertheimer (1880-1943), Wolfgang Köhler (1887-1967) e Kurt Koffka (1886-1940), os princípios gestálticos de agrupamento são usados especialmente pelo nosso cérebro quando a informação contida nos objectos é incompleta ou insuficiente. Estes princípios baseiam-se no conceito de inferência inconsciente de Helmholtz.

Nos nossos dias, os princípios gestálticos de agrupamento são muito usados em profissões como *Design* e *Arquitectura*. Tendem a descrever como é que as pessoas percebem os padrões organizados ou os conjuntos diferenciados dos elementos visuais, em vez dos elementos isolados, que constituem o todo, ou os próprios objectos.

Gestalt é uma palavra alemã que significa "configuração" ou "padrão." No âmbito da percepção visual, os psicólogos gestálticos postularam que agrupamos os elementos em padrões, de acordo com as seguintes regras (Shepard 1999, p. 32):

- Proximidade. Elementos que se encontram localizados próximos uns dos outros são agrupados como se fizessem parte do mesmo objecto (Figura 21).



Figura 21. Princípio gestáltico de proximidade. Adap. de Palmer et al. (2003, p. 312).

- **Similaridade.** Elementos que se encontram igualmente espaçados entre si tendem a agrupar-se com aqueles que aparentam ser seus similares. Ou seja, elementos que partilham características visuais como a forma, cor, textura, são percebidos como se pertencessem todas ao mesmo objecto (Figura 22).



Figura 22. Princípio gestáltico de similaridade. Adap. de Palmer et al. (2003, p. 312).

- **Simetria.** Apesar da distância, os elementos simétricos tendem a ser agrupados (Figura 23).

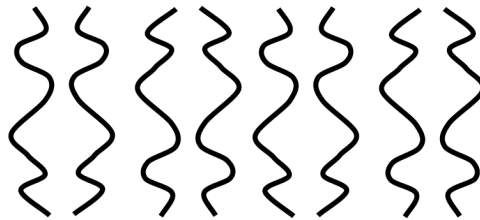


Figura 23. Princípio gestáltico de simetria. Adap. de Wagemans et al. (2012, p. 1180).

- **Boa continuação.** Elementos que aparentam seguir uma determinada direcção são percebidos como estando agrupados (Figura 24).

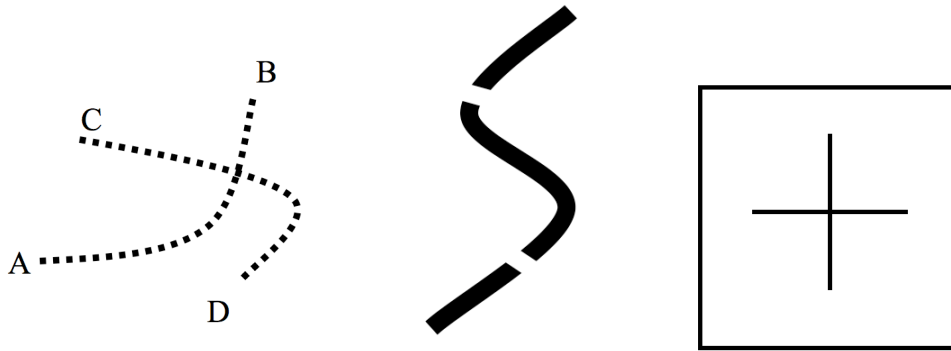


Figura 24. Princípio gestáltico de boa continuação. Adap. de Deutsch (1999, p. 300); Shepard (1999, p. 32).

Percepcionamos a imagem que se encontra mais à direita na Figura 24, como duas linhas contínuas que se cruzam, em vez de quatro linhas que se encontram no centro.

- Destino comum. Elementos que variam da mesma maneira ou que se movimentam à mesma velocidade e na mesma direcção tendem a ser agrupados (Figura 25).

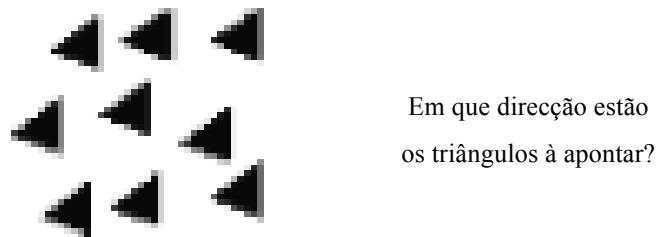


Figura 25. Princípio gestáltico de destino comum. Adap. de Palmer (2002, p. 212).

No que concerne ao sistema auditivo, os princípios gestálticos de agrupamento foram utilizados por Bregman (1990) para demonstrar o seu conceito de *Auditory Scene Analysis* (ASA). ASA, é o processo que o sistema auditivo utiliza para analisar e organizar os elementos dos sons provenientes de um determinado ambiente acústico, de maneira a atribuir-lhes significados perceptivos (Bregman 1993, p. 11). O sistema auditivo agrupa os elementos dos sons por proximidade temporal e por proximidade de frequência (Bregman 1990, p. 76). Para demonstrar estes princípios, Bregman criou uma sequência de seis sons puros tocados alternadamente (1996, exemplo 1 - *stream segregation in a cycle of six tones*). Três dos sons são de alta frequência (A1, A2 e A3) e

os outros três de baixa frequência (B1, B2 e B3). Ao tocar a sequência num tempo lento, os sons são percebidos alternadamente (A1-B1, A2-B2, A3-B3). Pelo contrário, quando tocada num tempo rápido (proximidade temporal) os sons deixam de ser percebidos de forma alternada, passando a escutar-se simultaneamente, dois grupos de sons contínuos com registos de altura distintos (segregação sonora). Um grupo constituído somente, por sons de alta frequência (A1-A2-A3) e outro, por sons de baixa frequência (B1-B2-B3) (proximidade de frequência). Esta experiência sugere, que numa sequência temporal rápida, as componentes sinusoidais dos sons que se encontram próximas em frequência tendem a ser agrupadas (Deutsch 1999, p. 313; Shepard 1999, p. 117; Bregman 1993, p. 21; Bregman 1990, p. 52).

Deutsch, através da sua ilusão sonora designada por *Octave Illusion* (1999, 1981, 1974) demonstrou que o sistema auditivo também agrupa os elementos dos sons, quer por similaridade de percepção de altura sonora, quer por proximidade de localização espacial. Para o efeito, criou dois sons puros com frequências separadas por um intervalo de oitava. Os sons, escutados através de auscultadores, foram apresentados simultaneamente e distribuídos de forma alternada por cada ouvido. Ao mesmo tempo que o ouvido direito recebe o som de alta frequência (a oitava superior) o ouvido esquerdo recebe o som de baixa frequência (a oitava inferior) e vice-versa. Em termos de percepção, a maior parte dos ouvintes ouve um único som que alterna entre as duas oitavas em sincronia com a mudança de localização entre os dois ouvidos. Ou seja, ouve um som agudo num ouvido alternando com um som grave no outro ouvido (Deutsch 1999, p. 332; Deutsch 1981, p. 100; Deutsch 1974, p. 307). Esta experiência sugere que sons puros, similares em termos de percepção de altura sonora e sons que aparentam ter a mesma localização espacial tendem a ser agrupados (Scharine e Letowski 2009, p. 581; Deutsch 1999, p. 334; Shepard 1999, p. 117; Bregman 1993, p. 26).

A *Octave Illusion*, em relação à experiência auditiva de Bregman, anteriormente mencionada, constitui um desafio auditivo maior, no sentido, de nela interagirem simultaneamente várias pistas auditivas. Para além do tempo, da frequência, temos também de lidar com a questão da localização espacial. Segundo Bregman (1990, p. 306) o principal resultado perceptivo alcançado revela o processo através do qual alguns ouvintes resolveram o “conflito” existente entre as várias pistas auditivas que se encontram presentes na ilusão. Ou seja: (1) a relação harmónica entre os dois sons puros (relação de oitava) sugere a presença de um único som, (2) a aparência de que o mesmo som alterna entre os dois ouvidos sugere que faz parte do mesmo evento sonoro e por

último (3) as localizações independentes dos sons puros sugerem a existência de dois sons separados.

Nesta ilusão, o princípio gestáltico de similaridade por percepção de altura sonora aparenta ser predominante sobre o princípio gestáltico de proximidade por localização espacial. De acordo com Bregman (1990, p. 306), se se eliminar a relação harmônica entre as duas frequências o efeito da ilusão desaparece.

Outro princípio gestáltico, designado por boa continuação, pode ser encontrado numa experiência auditiva elaborada por Warren (2008), designada por *Homophonic Continuity*. Neste caso concreto, estamos perante o princípio de boa continuação por percepção de intensidade sonora. Aliás, o termo *homophonic*, refere-se a sons que diferem unicamente em termos de intensidade sonora (Warren 2008, p. 152). Em termos gerais, a experiência consiste no seguinte: suponhamos que nos apresentam durante alguns segundos um som contínuo com uma determinada intensidade fixa. Repentinamente, o som torna-se mais intenso (durante um curto período de tempo) regressando rapidamente à sua intensidade inicial e continuando a ouvir-se por mais uns segundos antes de se extinguir. O som inicial é ouvido como contínuo, com a mesma percepção de intensidade sonora e por detrás de um som novo, que surge e desaparece subitamente (Bregman 1993, p. 15). O efeito aparente de continuação do som menos intenso mantém-se (Warren 2008, p. 151). Pelo contrário, se o aumento da intensidade for suave e progressivo a percepção da existência de dois sons não se concretiza (Bregman 1993, pp. 15-16). Esta última situação, exemplifica outro dos princípios gestálticos conhecido por destino comum. Por exemplo, se num som complexo, os parciais se movimentarem paralelamente, à mesma proporção e ao mesmo tempo, tendem a ser agrupados (Scharine e Letowski 2009, p. 581; Deutsch 1999, p. 300; Bregman 1993, p. 30; Nakajima *et al.*, 1988, pp. 1-5). Bregman (1996) demonstrou este princípio através do seu exemplo número 19, designado por “*Fusion based on common frequency modulation: illustration 1*”. O exemplo, baseia-se no seguinte: se as frequências das componentes de um som complexo permanecerem constantes ouviremos um único som. Se mantivermos as frequências de um determinado número de componentes constantes, enquanto aplicamos uma modulação de frequência às restantes componentes passaremos a ouvir dois sons distintos e separados (segregação sonora). Ouviremos, por um lado, um som estável e por outro lado, um som modulado. Por último, quando deixarmos de aplicar a modulação de frequência, as componentes moduladas regressam à sua posição inicial e passaremos a ouvir novamente, o som inicial.

Movimentos paralelos nas componentes dos sons promovem a sua fusão, enquanto que movimentos independentes provocam a sua segregação (Bregman 1993, p. 31).

Os princípios gestálticos, para além de desempenharem um papel importante no conceito de *Auditory Scene Analysis*, permitem-nos também, conhecer melhor os mecanismos de funcionamento do nosso sistema auditivo, quando confrontados, simultaneamente, com várias pistas auditivas ou efeitos sonoros provenientes, quer de ambientes acústicos naturais, quer de ambientes acústicos produzidos em situação de laboratório.

V ILUSÕES: GENERALIDADES

V.1 Definição de Ilusão

Nos dicionários generalistas (e.g., Priberam,⁷ Infopédia,⁸ *Oxford English Dictionary*,⁹ entre outros), encontramos várias definições para as ilusões. De uma forma sintética, todos eles definem as ilusões como algo que produz uma impressão falsa ou enganosa da realidade. Qualificam-nas como erros dos sentidos.

Num âmbito mais restrito, o da Psicologia Cognitiva, tanto Gregory (1997, p. 279), como Warren (1970, p. 30) definem as ilusões como erros da percepção. O primeiro autor considera ser importante limitar as ilusões às discrepâncias observadas pelos nossos sentidos, a partir de medições simples com réguas, relógios, fotómetros, audímetros, etc. O segundo autor, afirma ser importante estudar as ilusões, porque nos ajudam a isolar e a clarificar alguns dos mecanismos fundamentais que actuam sobre a percepção.

Por outro lado, Risset (2007, p. 6; 1992, p. 3) ao citar Purkinje, afirma que as ilusões são “erros dos sentidos e verdades da percepção”. As ilusões são consideradas como verdades da percepção, no sentido, em que nos ajudam a compreender os mecanismos de funcionamento da percepção humana.

Por vezes, confundem-se ilusões com alucinações. Para evitar, que tal situação aconteça faremos a seguinte distinção: as ilusões pressupõem a presença de estímulos. São sensações reais percebidas de maneira distorcida. Por exemplo: ouvir vozes no som de água a correr ou noutra fonte sonora qualquer. Pelo contrário, as alucinações são consideradas experiências sensoriais na ausência de qualquer estímulo (Scharine e Letowski 2009, p. 592). Por exemplo, há pessoas que ouvem vozes e sinos mesmo quando não está ninguém a falar, ou na ausência de sinos por perto.

⁷ <http://www.priberam.pt/dlpo/default.aspx?pal=ilusão>

⁸ <http://www.infopedia.pt/pesquisa-global/ilusão>

⁹ <http://www.oed.com/>

V.2 Classificação das Ilusões

Gregory (1968, p. 279) referindo-se às ilusões visuais afirma que existem ilusões puramente ópticas (i.e., físicas), onde a luz que emana dos objectos para os olhos é moldada, quer por fenómenos físicos de reflexão (espelhos), quer por fenómenos físicos de refacção (e.g., a imagem distorcida de um pau imerso em água, ou as miragens).

De acordo, com a sua perspectiva existem também ilusões sensoriais. Os órgãos dos sentidos (os olhos, os ouvidos, o nariz, a boca e a pele) podem ser perturbados transmitindo informações enganosas ao cérebro. Estes, podem ser perturbados por sobre-estimulação ou estimulação prolongada, como acontece no efeito de sobre-estimulação, que desencadeia a Ilusão Visual do Cubo de Necker (Figura 15), ou na observação prolongada de uma espiral em rotação (Figura 26).



Figura 26. Espiral. Adap. de Gregory (1968, p. 280).

Se colocarmos a espiral em rotação e se fixarmos o seu centro durante alguns segundos observamos que o seu tamanho se contrai ou se expande, de acordo, com o sentido da rotação. Para além disso, se pararmos a rotação da espiral acontece um movimento ilusório no sentido contrário ao da rotação original.

Para além destes dois tipos de ilusões, Gregory (1968, p. 280) acrescenta ainda um terceiro: as ilusões perceptivas. Estas, são originadas a partir da interpretação errada que o cérebro faz da informação sensorial recebida.

Em relação à classificação das ilusões sonoras devemos referir Féron (2006) e a sua Tese de Música e Musicologia com o título de: *“Des illusions auditives aux singularités du son et de la perception: l’impact de la psychoacoustique et des nouvelles*

technologies sur la création musicale au XXe siècle”. Féron, distingue dois tipos de fenómenos: os que resultam de propriedades físicas (fenómenos acústicos) e os que derivam da percepção (fenómenos auditivos). Este autor classifica as ilusões sonoras de acordo com os domínios perceptivos do som (altura, espaço, tempo, intensidade e timbre). Ou seja: (1) ilusões de percepção de altura sonora (e.g., ilusão de infinito e ilusão de intervalo), (2) ilusões espaciais (e.g., ilusão de movimento, ilusão de localização e ilusão de espaço específico ou ambiente acústico), (3) ilusões temporais (e.g., ilusão rítmica, ilusão de duração e ilusão de continuidade), (4) ilusões de intensidade, (5) ilusões referenciais e tímbricas (e.g., ilusão de identificação, ilusão de presença, ilusão de nome e ilusão semântica ou de transformações verbais). Por último, reagrupa numa última categoria as ilusões cinestésicas. Este tipo de ilusões, estão relacionadas com as interações sensoriais. Ou seja, os estímulos visuais e os estímulos sonoros influenciam-se mutuamente originando ilusões perceptivas.

Nesta dissertação, as ilusões sonoras que têm origem em causas naturais ou físicas, serão denominadas de ilusões físicas, por exemplo: ouvir vozes na água a correr, os ecos, o Efeito de *Doppler*, etc., e aquelas que têm origem na aplicação equívoca do nosso conhecimento, na interpretação de sinais sensoriais gerados a partir de processos artificiais e em contexto de laboratório, serão denominadas de ilusões cognitivas. Nesta dissertação e com o objectivo de direccionar o nosso trabalho decidimos concentrar a investigação exclusivamente no estudo das ilusões sonoras cognitivas. Dentro destas, iremos estudar somente as ilusões originadas a partir de fenómenos auditivos relacionados com o domínio das frequências, ou se preferirmos, com o espectro sonoro. Por exemplo, a Ilusão Sonora da Escala de Shepard (Shepard, 1964); o Paradoxo do Trítono (Deustch, 1974); o Glissando Perpétuo (Risset, 1969), etc.

De entre as várias ilusões cognitivas relacionadas com o domínio das frequências, decidimos estudar unicamente a Ilusão Sonora da Escala de Shepard, pelo facto de ter sido a ilusão que mais nos fascinou em termos sonoros e por ter revelado um grande potencial musical por explorar. Para além disso, contribuíram também para a nossa escolha, o facto da ilusão se apresentar sob a forma de uma escala musical e simultaneamente, por ser constituída por sons (i.e., o Som de Shepard) que apresentam a peculiaridade de serem ambíguos em termos de percepção de altura sonora, permitindo desta maneira, criar novas possibilidades de geração de som e de escuta musical.

V.3 Um Breve Estudo sobre os Processos de Geração de Ilusões

Um bom exemplo sobre o processo de construção de uma ilusão cognitiva encontra-se na descrição da Ilusão Visual de Ponzo (1913) (Figura 27).

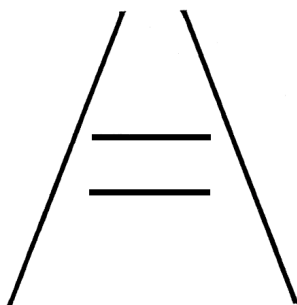


Figura 27. Ilusão Visual de Ponzo. Adap. de Gregory (1968, p. 284).

Na ilusão, a linha paralela superior parece estar expandida em relação à linha paralela inferior. No entanto, ambas possuem a mesma dimensão. O grafismo, representado na ilusão, baseia-se na fotografia dos Carris do Caminho de Ferro (Figura 28).



Figura 28. Fotografia dos Carris do Caminho de Ferro. Adap. de Gregory (1968, p. 282).

Mesmo retirando da fotografia os elementos que contribuem para a percepção tridimensional da imagem, o paradoxo de profundidade desaparece, mas as distorções de tamanho continuam presentes. Falamos em paradoxo de profundidade, porque apesar da fotografia conseguir representar a profundidade de uma paisagem, não deixa de ser um objecto plano bidimensional.

De uma maneira natural, directa e automática verificamos que com a distância, os carris (que são linhas paralelas) tendem a convergir à medida que se afastam, acabando

por se encontrar ou tocar num determinado ponto (perspectiva linear). No entanto, se caminharmos ao longo dos mesmos, com a intenção de alcançarmos o seu ponto de convergência constataremos que os carris nunca se chegam a tocar. Observamos também, que com a distância, as traves mais afastadas, onde assentam os carris, tornam-se menores em relação às traves mais próximas, que tendem a ser percebidas como maiores. Apesar, das distorções de tamanho permanecerem sabemos que na realidade, os carris são paralelos e que as traves são do mesmo tamanho.

No entanto, se desenharmos sobre a fotografia dos Carris do Caminho de Ferro dois rectângulos brancos da mesma dimensão, a ordem natural da nossa percepção é alterada (Figura 29).



Figura 29. Desenho de dois rectângulos brancos sobre a fotografia dos Carris do Caminho de Ferro. Adap. de Gregory (1968, p. 282).

O rectângulo superior que se encontra mais afastado deveria ser percebido como sendo o mais pequeno. No entanto, expande-se tornando-se maior do que o rectângulo inferior que se encontra mais próximo. Desta maneira, os elementos externos desenhados sobre a Figura 29, alteram a realidade. Através da perspectiva linear, o elemento externo mais distante expande-se, enquanto que o elemento que se encontra mais próximo se contrai.

As ilusões cognitivas resultam das alterações provocadas na ordem natural da realidade. São essas alterações que nos induzem a fazer interpretações erradas de dados sensoriais adquiridos ao longo da nossa aprendizagem e evolução.

Se neste momento, nos magoássemos fisicamente e víssemos surgir primeiro a cicatriz e só depois a ferida, como é que reagiríamos ou como é que o nosso cérebro interpretaria tal situação?

A alteração da ordem natural da realidade leva-nos a realizar leituras erradas de sinais sensoriais e induzem-nos a cometer erros de percepção. São os erros que estão na origem das ilusões cognitivas que nos permitem desvendar os mecanismos de funcionamento da percepção humana.

Poderemos simular e aplicar os pressupostos acima mencionados à percepção auditiva?

Bastará alterar a ordem natural da realidade para criar uma ilusão sonora? Por exemplo, se alternarmos a ordem natural dos acontecimentos gerados por um grito no cume de uma montanha, em que primeiro ouvimos os ecos e só depois o som que os originou, conseguiremos produzir uma ilusão sonora? Pressupomos que sim. Quando escutamos uma escala musical (e.g., uma escala dividida em 12 partes iguais) tocada de forma ascendente e em ciclo, estamos sempre à espera de ouvir um salto descendente na percepção de altura sonora entre a última nota (a sensível) e a primeira nota (a tónica) da respectiva escala. Este salto descendente, pela nossa experiência e ensino musicais é natural, é expectante, no fundo, é uma realidade adquirida. No entanto, se por qualquer processo artificial, o salto descendente for anulado, a nossa expectativa inicial e a realidade que dávamos por adquirida, são questionadas. Este paradigma, foi concretizado através da Ilusão Sonora da Escala de Shepard. A ilusão (que irá ser amplamente abordada na parte II, desta dissertação) altera a realidade, ou seja, altera a percepção auditiva da respectiva escala musical. Por isso, nos surpreende e cria perplexidade, no sentido em que algo, enraizado e dado como adquirido, deixa de o ser sempre que a escutamos.

V.4 Mecanismos de Actuação das Ilusões na Percepção Humana

Em sentido lato, as ilusões cognitivas actuam sobre a percepção gerando ambiguidades ou erros sistemáticos de percepção, que nos induzem a fazer interpretações erradas dos sinais sensoriais. Mencionamos erros sistemáticos, porque apesar de nos expormos vezes sem conta à mesma ilusão, em que os factores surpresa e novidade já são

praticamente nulos, continuamos sempre a perceber os mesmos efeitos gerados pela respectiva ilusão.

Factores como a sobre-estimulação ou a estimulação prolongada (que desencadeiam a Ilusão Visual do Cubo de Necker, a *Loop Words* de Warren, *afterimages*, etc.), a perspectiva linear (que contribui para a Ilusão Visual de Ponzo), os princípios de inferência inconsciente (que se encontram na Ilusão Visual das pernas do elefante de Shepard e no fenómeno psicoacústico da *missing fundamental*), bem como a tendência natural de tentarmos encontrar padrões no meio ambiente que nos rodeia (princípios geltálticos de agrupamento demonstrados através do conceito de *Auditory Scene Analysis*) apontam ser os principais factores (cada um deles concorrente à sua maneira), que se encontram na origem das ilusões cognitivas apresentadas.

A seguir, vamos abordar a Ilusão Sonora da Escala de Shepard, no que respeita à sua caracterização, percepção auditiva, trabalhos relacionados, implementação em síntese de som digital e análise espectral.

PARTE II - ILUSÃO SONORA DA ESCALA DE SHEPARD

VI ILUSÃO SONORA DA ESCALA DE SHEPARD

Shepard construiu uma ilusão sonora designada por Ilusão Sonora da Escala de Shepard (1964) (*The Shepard Scale Illusion*).¹⁰ A ilusão é uma escala musical dividida em 12 partes iguais. Quando escutada pode gerar a ilusão sonora de subir ou descer infinitamente, de acordo, se é tocada de forma ascendente ou descendente, ao longo de uma espiral, como acontece no plano visual com a Escada Interminável de Escher (1960) (Fugiel 2011, p. 415; Shepard 1964, p. 2349) (Figura 30).

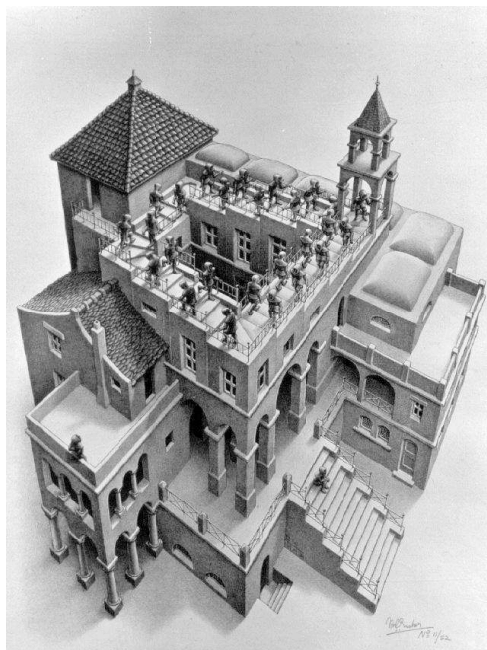


Figura 30. Litografia da Ilusão Visual da Escada Interminável de Escher. Na parte superior da imagem é possível observar os monges a subir ou descer as escadas ao longo de um ciclo interminável. Adap. de Hofstadter (1979, p. 20).

¹⁰ O som original da Ilusão Sonora da Escala de Shepard pode ser descarregado em http://asawebdev.devcloud.acquia-sites.com/listen_to_sounds/08_03_10_demo27 ou ouvido no DVD em anexo, no directório sons, com o nome *01.circularity_in_pitch_judgment.wav*

Para um visionamento pormenorizado da escada que se encontra na parte superior da litografia apresentamos a Figura 31.

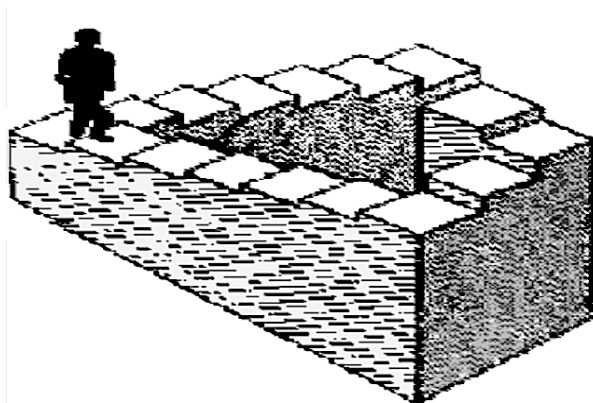


Figura 31. Imagem produzida a partir da parte superior da litografia de Escher. Por mais que a figura humana suba ou desça as escadas permanecerá enclausurada num ciclo interminável.¹¹

De acordo, com Hofstadter (1979, p. 18), a Ilusão Visual da Escada Interminável de Escher representa um ciclo estranho (*strange loop*). Um ciclo estranho acontece, quando nos deslocamos ao longo das etapas de um sistema qualquer (visual, musical, matemático, etc.) e de repente, de forma imprevisível damos connosco de volta ao ponto de partida.

A Ilusão Sonora da Escala de Shepard também é semelhante à Ilusão Visual de um Poste de Barbeiro em movimento (Loy 2006, p. 165; Shepard 1999, p. 158) (Figura 32).

¹¹ Adap. de <http://www.scientificpsychic.com/graphics/index.html>



Figura 32. O Poste de Barbeiro possui faixas vermelhas e azuis dispostas na diagonal sobre um fundo branco. Quando colocadas a girar e observadas de um determinado ângulo parecem mover-se para cima ou para baixo, ao longo de um movimento contínuo no sentido da rotação.¹²

Em termos gerais, Shepard concebeu a ilusão sonora anulando a dimensão vertical dos sons através da construção do Som de Shepard (*Shepard tone*).

VI.1 Som de Shepard

O Som de Shepard é um som complexo constituído por dez componentes separadas por intervalos de oitava e às quais foi aplicada, tendo em consideração uma escala logarítmica de frequência, uma envolvente espectral em forma de sino ou de Curva Gaussiana (Figura 33).

¹² Ver o Poste de Barbeiro em movimento no DVD em anexo, no directório vídeos, com o nome *01.poste_barbeiro.mov*

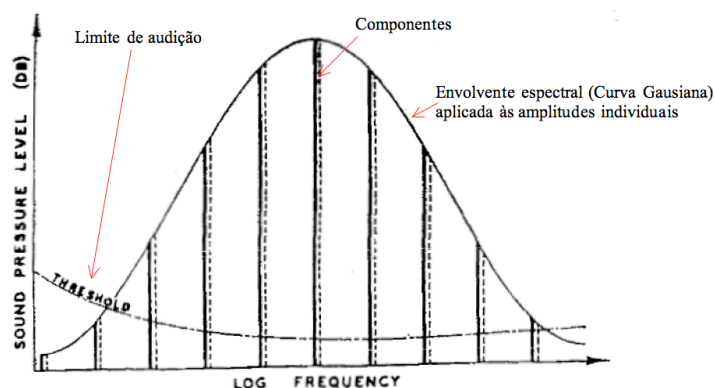


Figura 33. Som de Shepard. Na imagem estão representadas as intensidades (dB) de dez componentes separadas por intervalos de oitava simbolizadas pelas linhas verticais contínuas. As linhas verticais a tracejado correspondem a um deslocamento ascendente nas frequências de todas as componentes. As componentes das extremidades possuem sempre as amplitudes mais baixas e as componentes intermédias as amplitudes mais altas. Adap. de Shepard (1964, p. 2347).

Shepard construiu o som através de um programa desenvolvido por Max Mathews (1926-2011) (Deutsch 2010, p. 8; Shepard 1999, p. 158).¹³ Os processos de síntese de som digital foram realizados nos Laboratórios da Telefónica Bell, em New Jersey (EUA). A síntese de som foi realizada através de um computador IBM 7094 (Shepard 1964, p. 2347).

O Som de Shepard é um som peculiar, no sentido, em que se torna ambíguo em termos de altura espectral, mas não se torna ambíguo em termos de altura tonal (Deutsch 2008, p. 589; Shepard 1999, p. 158). Provavelmente, somos capazes de identificar a nota musical do som (i.e., a altura tonal), mas teremos muita dificuldade em determinar a oitava a que o som pertence (i.e., a altura espectral) (Fugiel 2011, p. 415).

Podemos afirmar, que o Som de Shepard não pertence a nenhuma oitava em particular, mas pertence a todas ao mesmo tempo, pelo facto, de todas elas (i.e., as dez oitavas que constituem um Som de Shepard) se encontrarem representadas simultaneamente, no mesmo som.

Shepard (1964, pp. 2347-2348) utilizou equações matemáticas para determinar, quer as frequências, quer as amplitudes individuais das componentes que constituem os sons da respectiva ilusão.

¹³ Para uma descrição detalhada do programa consultar o Artigo de Max Mathews intitulado “The digital computer as a musical instrument”, *Science*, 142, 1963, pp. 553-557.

VI.1.1 Determinação Manual das Frequências Individuais

Para determinar manualmente as frequências individuais utiliza-se a Equação 2 (Shepard 1964, p. 2347).

$$f(t, c) = f_{\min} \cdot 2^{[(c-1)t_{\max} + t - 1] / t_{\max}}, \quad (2)$$

Onde:

- $f_{\min} = 4.863$ Hz, representa a frequência da primeira componente (i.e., a frequência fundamental do primeiro som complexo da escala);
- $t_{\max} = 12$, representa o número de sons complexos que constituem a escala;
- t representa o som complexo;
- c representa a componente do som complexo.

Os valores mencionados são constantes e foram fornecidos pelo autor (Shepard 1964, p. 2348).

Para obter as frequências das componentes para o primeiro som da escala atribui-se à variável t o valor um e aplica-se de forma progressiva os valores de um a dez à variável c .

Para obter as frequências dos restantes sons aplica-se o procedimento anterior subindo progressivamente um valor à variável t , e.g., $t = 1$, $t = 2$, $t = 3$, e assim sucessivamente, até $t = 12$.

VI.1.2 Determinação Manual das Amplitudes Individuais

Para determinar manualmente as amplitudes individuais calcula-se em primeiro lugar o valor de θ através da Equação 3 (Shepard 1964, p. 2347).

$$\theta(t, c) = 2\pi[(c - 1) * t_{\max} + t - 1] / [t_{\max} * c_{\max}], \quad (3)$$

Onde:

- $t_{\max} = 12$, representa o número de sons complexos que constituem a escala;
- $c_{\max} = 10$, representa o número máximo de componentes de cada som complexo.

Os valores apresentados são constantes e foram fornecidos pelo autor (Shepard 1964, p. 2348).

Para obter os valores de θ para as componentes do primeiro som da escala atribui-se à variável t o valor um e aplica-se de forma progressiva os valores de um a dez à variável c .

Para obter os valores de θ dos restantes sons aplica-se o procedimento anterior subindo progressivamente, um valor à variável t , e.g., $t = 1$, $t = 2$, $t = 3$, e assim sucessivamente, até $t = 12$.

De seguida, aplica-se a Equação 4 (Shepard 1964, p. 2347).

$$L(t, c) = L_{\min} + [L_{\max} - L_{\min}] * [1 - \cos \theta(t, c)] / 2 \quad (4)$$

Onde:

- $L_{\min} = 22$ dB, representa o limite mínimo do nível de pressão sonora das amplitudes individuais;
- $L_{\max} = 56$ dB, representa o limite máximo do nível de pressão sonora das amplitudes individuais.

Os valores mencionados em dB são constantes e foram fornecidos pelo autor (Shepard 1964, p. 2348).

VI.2 Escala de Shepard

Shepard temperou cada um dos 12 sons complexos (i.e., atribuiu-lhes uma altura sonora definida, respeitando os parâmetros de afinação utilizados na música ocidental, em que o A4 possui o valor de 440 Hz) de maneira, a obter uma escala musical dividida em 12 partes iguais. Ao anular a dimensão vertical dos sons, demonstrou por um lado, que a altura espectral e a altura tonal são dois atributos independentes da percepção auditiva (Deutsch 1986, p. 1346; Shepard 1964, p. 2347). Por outro lado, demonstrou que a percepção da altura dos sons e as frequências são duas entidades sonoras diferentes (Pollack 1978, p. 202).

O *American National Standard Institute* (ANSI) define a percepção de altura dos sons como sendo o atributo da percepção auditiva que nos permite ordenar os sons numa escala que se estende desde os sons mais agudos até aos sons mais graves (Deutsch 2008, p. 589; Risset 1971, p. 613).

No entanto, a percepção de altura dos sons não é um atributo exclusivamente unidireccional (Deutsch 1992, p. 391). Em termos particulares, os músicos e os compositores sabem que existe uma similaridade muito forte na percepção auditiva de sons relacionados por intervalos de oitava (Deutsch 2008, p. 589). As escalas musicais baseadas nesta similaridade atribuem aos sons separados por intervalos de oitavas os mesmos nomes, por exemplo: C, D, E, e assim sucessivamente (Deutsch 1972, p. 1).

A percepção de altura de sons complexos, cujas componentes se encontram separadas por intervalos de oitava, pode ser representada através de um modelo bidimensional (Burns 1981, p. 467; Deutsch 1986, p. 1346; Loy 2006, p. 164; Shepard 1999, p. 157; Risset 1971, p. 613).

Este modelo, constituído por uma hélice apresenta uma dimensão rectilínea ou vertical – altura espectral, que representa a posição do som ao longo de um contínuo e pela dimensão circular ou horizontal – altura tonal, que representa a posição do som dentro da oitava (Burns, 1981, p. 467; Deutsch 2008, p. 589; Risset 1977, p. 2; Risset 1971, p. 613; Shepard 1964, p. 2346; Ueda e Ohgushi 1987, p. 1193) (Figura 34).

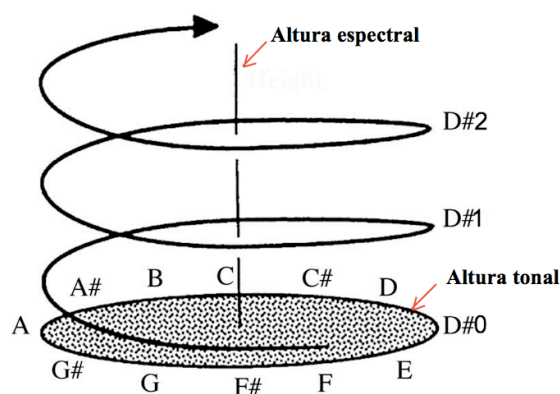


Figura 34. Representação bidimensional da percepção de altura de sons complexos, cujas componentes se encontram separadas por intervalos de oitava. Na Figura, encontramos representada a dimensão vertical dos sons (altura espectral) e a dimensão circular (altura tonal). Adap. de Deutsch (2008, p. 590).

Na Figura 34, a altura espectral está representada no eixo vertical e a altura tonal na escala circular que constitui a base da hélice (Shepard 1964, p. 2346). Como a hélice realiza uma volta completa por cada oitava permite que os sons, que mantêm relações de equivalência de oitava, se encontrem espacialmente mais próximos, como se pode comprovar pela notas D#0, D#1, D#2 (Deutsch 1986, p. 1346; Shepard 1964, p. 2346).

Na notação musical anglo-saxônica, esta representação é usada da seguinte maneira: as notas musicais são representadas por letras maiúsculas que designam o nome e a posição que ocupam dentro da oitava e por um número que designa a oitava em que as notas se encontram inseridas (Deutsch 2008, p. 589). Por exemplo: D2, D3, D4, D5, etc., designam notas musicais com a mesma altura tonal que se encontram em oitavas diferentes e C3, F#3, A3, B3, etc., designam notas que se encontram dentro da mesma oitava, mas com altura tonal diferente (Deutsch 2008, p. 589).

Numa progressão melódica descendente por intervalos ascendentes de sétima maior, por exemplo C2-B2, Bb2-A3, Ab3-G4 e Gb4-F5, ouviremos provavelmente, um efeito sonoro ambíguo, devido ao facto, de os sons se encontrarem simultaneamente, distantes em termos de altura espectral (i.e., 11 meios-tons de C2 a B2) e muito próximos em termos de altura tonal (i.e., meio-tom de B2 a C2). A sequência, que tende a ser naturalmente escutada como ascendente (em termos de percepção de altura sonora) pode, com um pouco mais de persistência auditiva, ser percebida como uma progressão

melódica descendente por meio-tom cromático descendente (Loy 2006, p. 165; Shepard 1964, p. 2346).¹⁴

Qualquer um destes atributos da percepção de altura de sons complexos (que se encontram correlacionados) pode ser isolado através de processos de síntese de som digital.

Um processo através do qual se pode eliminar a altura tonal, e simultaneamente, conservar a altura espectral, passa por filtrar ruído branco através de um filtro Passa-Banda (Shepard 1999, p. 160; Shepard 1964, p. 2351). Neste exemplo, a frequência central do filtro determinará a percepção de altura sonora.¹⁵

O fenómeno inverso também é possível. Através do Som de Shepard consegue-se eliminar a altura espectral, e simultaneamente, preservar a altura tonal. Provavelmente, conseguiremos identificar o nome da nota musical do som (a altura tonal), mas teremos muita dificuldade em determinar a oitava a que pertence (a altura espectral) (Fugiel 2011, p. 415).

Deutsch demonstrou que a altura tonal é uma dimensão essencial na percepção de altura dos sons. Através da transposição aleatória de oitavas de algumas notas musicais de uma melodia popular norte-americana (mantendo simultaneamente a altura tonal) constatou que os ouvintes (independentemente da sua experiência e treino musical) não foram capazes de identificar o tema musical popular escutado (Deutsch 1972, pp. 411-412). A Figura 35, exemplifica essa demonstração realizada através de outra melodia.

¹⁴ O exemplo pode ser consultado no DVD em anexo, no directório *patches*, com o nome *01.efeito_sonoro_ambiguo.pd*

¹⁵ O exemplo pode ser consultado no DVD em anexo, no directório *patches*, com o nome *02.tone_height.pd*

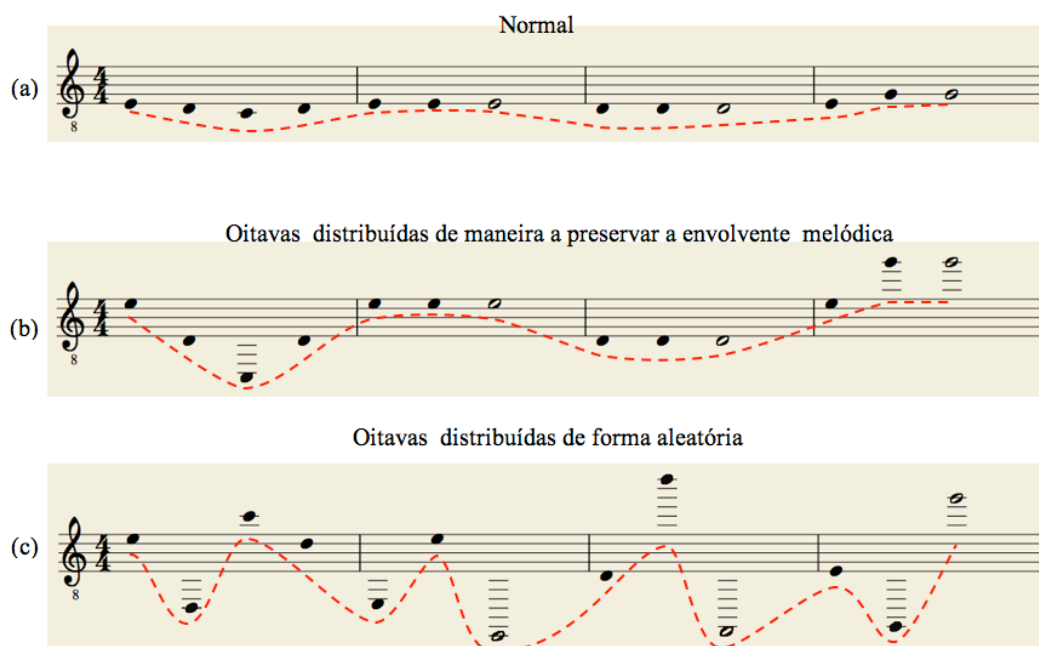


Figura 35. Quando as oitavas de uma melodia conhecida são distribuídas de forma aleatória (c)¹⁶ torna-se bastante difícil reconhecer a melodia original (a).¹⁷ Pelo contrário, quando se distribuem de maneira a preservar a sua envolvente melódica (b)¹⁸ torna-se mais fácil o reconhecimento da melodia original. Adap. de Shepard (1999, p. 161).

Na Ilusão Sonora da Escala de Shepard cada som ocupa uma posição fixa e representa sempre a mesma altura sonora ao longo de qualquer oitava. Por exemplo, a nota musical C representa, numa guitarra ou noutro instrumento qualquer, todas as classes de altura com esse nome. A dimensão circular ou horizontal dos sons permite representá-los visualmente como pontos igualmente espaçados à volta de um círculo (Shepard 1964, p. 2348) (Figura 36).

¹⁶ O exemplo pode ser escutado no DVD em anexo, no directório sons, com o nome *05.melodia_aleatoria.wav*

¹⁷ O exemplo pode ser escutado no DVD em anexo, no directório sons, com o nome *03.melodia_original.wav*

¹⁸ O exemplo pode ser escutado no DVD em anexo, no directório sons, com o nome *04.melodia_preservada.wav*

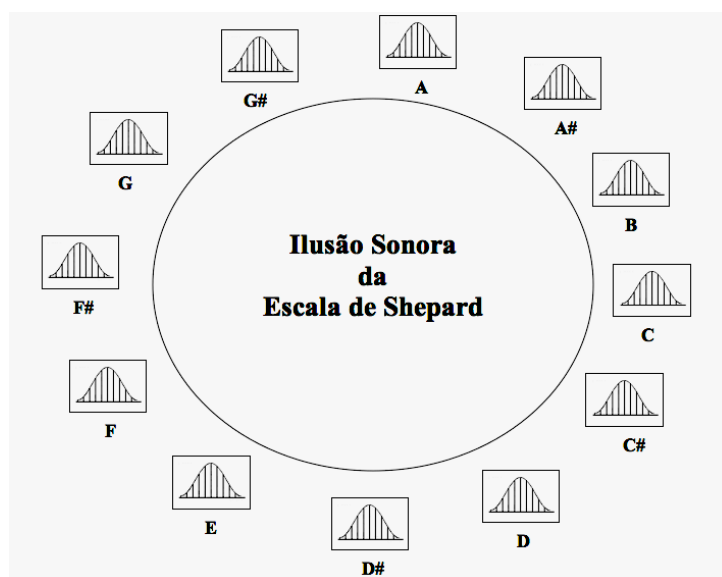


Figura 36. Representação da escala musical da Ilusão Sonora da Escala de Shepard. A imagem representa a relação circular ou horizontal dos sons complexos, ao longo da respectiva ilusão. A ilusão é uma escala musical dividida em 12 partes iguais. Adap. de Deutsch (1987, p. 563).

VI.2.1 Percepção Auditiva da Ilusão Sonora da Escala de Shepard

Na ilusão, os sons complexos tem a duração de 120 milissegundos e encontram-se separados por uma pausa de 840 milissegundos de duração. A fita áudio, que contém a ilusão, foi tocada para mais de 60 ouvintes (Shepard 1964, p. 2347-2348). Todos os ouvintes afirmaram, que a sequência progredia linearmente em altura, de forma ascendente ou descendente dependendo, se a fita era tocada para a frente (*forward*) ou para trás (*rewind*) (Shepard 1964, p. 2348). Ao ser tocada para a frente, cada som era geralmente ouvido como mais agudo do que o som anterior (Shepard 1964, pp. 2348-2349). Alguns ouvintes intrigaram-se com o facto de que os sons (que claramente tinham estado a subir durante algum tempo) não pareciam tornar-se mais agudos (Shepard 1964, p. 2349). Outros, no entanto, não perceberam esta estagnação na altura dos sons. Estes ouvintes ficaram surpreendidos ao saber que a sequência era cíclica em vez de linear e que regressava com precisão ao som inicial (Shepard 1964, p. 2349).

Para além disso, outros ouvintes (principalmente aqueles que possuíam treino musical) detectaram a ocorrência de um salto na escala, no preciso momento em que esta regressava à tónica sobretudo, quando lhes foi apresentada a ilusão sem o atraso de 840 milissegundos que ocorre entre os sons (Shepard 1964, p. 2349).

Apesar do Som Shepard ser constituído por dez componentes (em que cada componente corresponde a um som puro ou sinusoidal) os sujeitos não foram capazes de os isolar auditivamente. Descreveram que ouviam um único som semelhante ao som de um órgão (Shepard 1964, p. 2348). Os movimentos paralelos das frequências das componentes promovem a sua fusão (Bregman 1993, p. 31) e por outro lado, a relação harmónica entre as frequências das componentes (i.e., a relação de oitava) sugere a presença de um único som (Bregman 1990, p. 306).

No plano visual, os ouvintes compararam a Ilusão Sonora da Escala de Shepard com a Ilusão Visual da Escada Interminável de Escher (Shepard 1964, p. 2349).

Ao pedir para identificar qual a sensação de direcção produzida por dois sons complexos, cujas direcções tonais estavam em relação ascendente, ou seja, A-B, C-E, D-A, Shepard constatou que (1) intervalos inferiores a meia-oitava (trítano) produziam uma sensação de subida na percepção de altura sonora, (2) intervalos entre meia-oitava e uma oitava produziam uma sensação de descida, uma vez que as suas fundamentais estão mais próximas no sentido descendente e (3) intervalos separados por uma oitava, não produziam qualquer alteração (Pollack 1978, p. 202).

VI.2.2 Factores que Contribuem para a Realização da Ilusão Sonora da Escala de Shepard

Existem vários factores que contribuem de diferentes formas para o sucesso auditivo da ilusão. Referimo-nos concretamente:

1. À utilização de uma envolvente espectral em forma de sino ou de Curva Gaussiana. Este tipo de envolvente permite realizar uma transição suave entre as componentes, realçar as componentes da região intermédia, e simultaneamente, tornar inaudível as componentes das extremidades (Pollack 1978, p. 202).
2. Ao facto, de em todos os sons complexos, à medida que a frequência da primeira componente aumenta, a intensidade das frequências das componentes superiores vai diminuindo gradualmente. Simultaneamente, a intensidade das frequências das componentes inferiores vai aumentando de forma progressiva. Assim, após se completar um ciclo (i.e., uma oitava) o estímulo torna-se idêntico ao inicial. Consequentemente, o salto descendente, na percepção de altura sonora entre a última

e a primeira nota musical da escala é atenuado, criando-se desta maneira, uma sensação de subida interminável na percepção de altura sonora (Ohgushi e Ueda 1987, p. 1193).¹⁹

3. À utilização de sons complexos. Os sons complexos (i.e., sons constituídos por mais do que uma componente) são timbricamente mais ricos do que os sons puros (i.e., sons constituídos por uma única componente). Consequentemente, podem gerar ambiguidades na percepção da altura sonora. Ao contrário dos sons complexos, os sons puros constituídos por uma única frequência têm uma localização bem definida na membrana basilar e por essa razão (pelo menos a partir do registo médio até próximo do agudo) não geram grandes ambiguidades na percepção da altura sonora. Nos sons complexos, as altas frequências das componentes interagem sobre o mesmo local da membrana basilar, fazendo com que esta se torne menos selectiva, ou seja, menos eficiente na tarefa de isolar as altas frequências das respectivas componentes (Plack 2004, pp. 14-15) originando por isso, uma maior indefinição na percepção da altura sonora. Sons com uma ou mais componentes isoladas tendem a gerar uma percepção de altura sonora bem definida, enquanto que sons constituídos por componentes não isoladas geram indefinições na percepção da altura sonora (Cheveigné 2010, p. 81).
4. A frequência fundamental é omissa e os sons podem ser ambíguos em termos de altura sonora (Pierce 1999, p. 184). Diríamos, que a frequência fundamental não é omissa, porque faz parte da constituição dos sons. O que acontece, é que as frequências das primeiras componentes de cada som se encontram abaixo do limiar da audição humana, e por isso, tornam-se menos audíveis. Referimo-nos a valores de frequências que variam entre os 4.863 Hz (i.e., a frequência da primeira componente do primeiro som) e os 9.18012 Hz (i.e., a frequência da primeira componente do décimo segundo som).
5. A factores que podem ser explicados através dos princípios gestálticos de agrupamento. Nomeadamente:

5.1. Destino comum. Os Sons de Shepard deslocam-se sempre no mesmo sentido ao longo da escala dividida em 12 partes iguais. As alterações ocorridas nas frequências das componentes são percebidas como alterações gerais quando as direcções e os

¹⁹ Apesar de o estímulo se tornar idêntico ao inicial, este não quebra a sensação de perpetuidade sonora da ilusão, porque estamos a lidar com sons ambíguos em termos de percepção de altura sonora (Fugiel 2011, p. 415; Deutsch 2008, p. 589; Shepard 1999, p. 158).

graus dessas mudanças são idênticas para todos elas (Shepard 1993, p. 31; Nakajima *et al.*, 1988, pp. 4-5).

5.2. Proximidade. Os resultados da experiência de Shepard apelam à influência do princípio de proximidade (Shepard 1964, p. 2352). Quando dois sons complexos são apresentados sucessivamente as frequências das componentes mais próximas de ambos os sons (logaritmicamente determinadas) tornam-se auditivamente ligadas (Nakajima *et al.*, 1988, pp. 4-5).

5.3. Similaridade. Quando as frequências das componentes estabelecem relações harmônicas de equivalência de oitava, tendem a ser agrupadas (Bregman 1990, p. 306; Nakajima *et al.*, 1988, p. 19).

5.4. Boa continuação. Elementos que aparentam seguir uma determinada direção são percebidos como estando agrupados.

VI.2.3 Som de Shepard e Ilusão Sonora da Escala de Shepard Revisitados

Em 1968, Risset elaborou uma variante da Ilusão Sonora da Escala de Shepard através da construção de um glissando sonoro que dá a sensação de subir ou descer perpetuamente, conforme é tocado de forma ascendente ou descendente (Risset 2009, p. 27; Risset 2007, p. 6; Risset 1998, p. 8; Risset 1994, p. 258; Risset 1992, p. 7; Risset 1977, p.7; Risset 1971, p. 614). Esta variante é conhecida por Glissando Perpétuo. Esta ilusão sonora de Risset encontra-se em algumas das suas obras musicais sobre suporte fixo integralmente compostas através de processos de síntese de som digital (Risset, 1994, p. 259). Por exemplo:

- *Computer Suite for Little Boy* (1968), segundo andamento (*Fall*), entre o minuto 0' e 1'10;
- *Mutations* (1969), entre o minuto 5'44 - 6'20 e entre o minuto 6'35 - 7'39.

Shepard considera serem essenciais na sua ilusão sonora os 840 milissegundos de atraso na sequência, ou no retardamento da sua progressão, que resulta desses atrasos (Shepard 1964, p. 2349).

No entanto, ao contrário de Shepard, que atribui grande importância às pausas entre os sons na realização da ilusão, Risset demonstrou não serem essenciais, quando os dispensou ao produzir o seu Glissando Perpétuo (Risset 2009, p. 27).

Pollack (1978, p. 202) realizou várias experiências na tentativa de identificar as principais características da demonstração de Shepard (1964). A tarefa dos ouvintes consistia em ouvir dois sons complexos e indicar se o segundo som era mais alto ou mais baixo em termos de percepção de altura sonora, ou se era mais brilhante, ou menos brilhante em termos de timbre. Ao contrário de Shepard, que ajustou a envolvente espectral em forma de sino às amplitudes individuais, através de equações matemáticas (Equação 3 e Equação 4), Pollack aplicou às amplitudes valores fixos e ponderados de forma logarítmica. Tendo obtido os mesmos resultados que Shepard, concluiu que a envolvente em forma de sino não é um factor primordial na demonstração e que a condição fundamental reside sobretudo, no número mínimo de componentes que um som complexo deve ter, provavelmente três componentes distribuídas e relacionadas num intervalo espectral entre os 100 Hz e 10000 Hz (Pollack 1978, p. 203).

Burns (1981, pp. 469-470) demonstrou que a Ilusão Sonora da Escala de Shepard pode ser produzida através de sons complexos inarmónicos, ou seja, sons cujas componentes não são múltiplas ou sub-múltiplas inteiras da frequência fundamental. Concluiu que (1) não existem diferenças significativas nas respostas entre a percepção de componentes que se encontram separadas por intervalos de oitava e por aquelas que se encontram separadas por relações intervalares inarmónicas, (2) não há uma correlação directa entre as respostas obtidas e o grau de treino musical dos sujeitos que se submeteram à experiência, (3) não encontrou indicações de que os sujeitos com maior treino musical tenham achado a percepção da ilusão sonora resultante das componentes separadas por intervalos de oitava (Ilusão Sonora da Escala de Shepard) mais evidente do que a mesma ilusão resultante de componentes separadas por relações intervalares inarmónicas e (4) devido à elevada correlação entre o treino musical e o grau de percepção na equivalência de oitava encontrado noutros estudos, a falta dessa correlação no trabalho de Burns sugere que a equivalência de oitava não é um factor preponderante, como Shepard inicialmente sugeriu.

Schroeder (1986, p. 188) gerou um paradoxo sonoro através de um fractal (i.e., uma função matemática). Sintetizou uma forma de onda constituída por oito frequências separadas por intervalos de oitava, que ao ser tocada no dobro de velocidade é percebida meio-tom abaixo em vez de uma oitava superior, como esperado.

Ueda e Ohgushi (1987) investigaram a possibilidade de representar os atributos da percepção de altura dos sons (i.e., a altura espectral e a altura tonal) sob a forma de um modelo em espiral tridimensional (Ueda e Ohgushi 1987, p. 1199). Ao contrário dos modelos teóricos anteriormente propostos, o seu modelo de representação baseou-se em experiências realizadas a partir da própria percepção de altura dos sons (Ueda e Ohgushi 1987, p. 1194). Construíram 18 Sons de Shepard modificados, ou seja, as frequências das componentes de cada som complexo, estão separadas por intervalos de um tom e cada som apresenta uma envolvente espectral diferente (i.e., o pico máximo da envolvente varia ao longo dos sons). De seguida, distribuíram os 18 sons por seis subgrupos, sendo cada um deles constituído por três sons. Por último, os seis subgrupos foram ordenados de forma aleatória e gravados em fita. A tarefa dos ouvintes consistia em indicar qual dos dois pares apresentados era percebido como sendo mais agudo em termos de percepção de altura sonora. Observaram que as diferenças individuais na percepção da altura dos sons resultavam da importância que os ouvintes atribuíam a cada um dos atributos da percepção, ou seja, à altura espectral ou à altura tonal (Ueda e Ohgushi 1987, p. 1199).

Nakajima *et al.* (1988) realizaram um estudo sobre a percepção de altura sonora de sons complexos através de tríades maiores. Numa das suas experiências (Nakajima *et al.*, 1988, p. 10) construíram dez sons complexos semelhantes aos Sons de Shepard. Cada som complexo era constituído por nove componentes. As componentes encontravam-se temperadas de maneira a formarem três tríades maiores (e.g., D3-F#4-A4, D4-F#5-A5 e D5-F#6-A6). Este tipo de sons complexos foram designados pelos autores de *chordal tones*. Os *chordal tones* foram apresentados aos pares (padrão sonoro). Por exemplo: um *chordal tone* de C maior (C4-E4-G4, C5-E5-G5 e C6-E6-G6) e outro de E maior (E4-G#4-B4, E5-G#5-B5 e E6-G#6-B6) de maneira, que o par partilhasse entre si algumas notas, ou pelo menos, notas que se encontrassem próximas em termos de altura tonal. Neste caso concreto, ambos os *chordal tones* partilham um par de notas (E-E) e dois pares de notas próximas (C-B) e (G-G#). A tarefa dos ouvintes consistia em julgar, em termos de percepção de altura sonora, se o segundo *chordal tone* era mais alto ou mais baixo do que o primeiro (Nakajima *et al.*, 1988, p. 8-11), ou se detectavam algum movimento na percepção de altura sonora (*dynamic pitch*) (Nakajima *et al.*, 1988, p. 5). Os *chordal tones*, que constituem cada um dos pares, eram tocados sequencialmente e separados por uma pausa de 600 milissegundos de duração. Em termos gerais, Nakajima *et al.*, (1988, pp. 18-19) concluíram que (1) as componentes (i.e., as notas musicais) que sobem ou

descem juntas podem propiciar a detecção de movimentos na percepção de altura sonora, mesmo quando estas não mantêm intervalos musicais constantes, (2) é possível detectar um ligeiro movimento na percepção de altura sonora, mesmo quando algumas componentes não se movem, (3) quando algumas componentes se deslocam em conjunto, na mesma direcção, por graus regulares e constantes podem originar movimentos na percepção de altura sonora, (4) quando as componentes se movimentam em conjunto, determinados ouvintes conseguem perceber as diferenças de meio-tom, enquanto que outros não, (5) que a periodicidade espectral das componentes (logaritmicamente determinadas) promove a percepção circular de altura dos sons (*circularity in pitch*) e por último (6) a percepção circular da altura dos Sons de Shepard (*pitch circularity of Shepard tones*) em que os sons se movimentam de meio-tom em meio-tom representa um fenómeno diferente da percepção circular da altura tonal (*circularity of tone chroma*) em que os sons se movimentam por intervalos superiores a meio-tom. Para além disso, para Nakajima *et al.*, (1988, p. 3) os sons de Burns não representam o fenómeno psicoacústico da percepção circular da altura dos sons (*circularity in pitch*), porque o ciclo de 12 sons nunca regressa com precisão ao ponto de partida.

Deutsch (1992, p. 391) apresentou um novo paradoxo sonoro conhecido pelo Paradoxo do Trítano. Com este paradoxo, a autora pretendeu avaliar a influência que a posição dos sons na oitava tem na percepção da altura sonora. Como no padrão sonoro (trítano) os sons se encontram em posições opostas na escala circular, o factor de proximidade não pode ser usado para explicar a percepção de altura dos sons. Assim, concluiu que (1) dois sons complexos que se encontram separados por meia oitava podem ser percebidos auditivamente como ascendentes numa tonalidade e descendentes quando tocados noutra tonalidade, (2) dentro de uma única oitava, determinados sujeitos percebem o intervalo como ascendente, enquanto que outros o ouvem como descendente, (3) a percepção do Paradoxo do Trítano poderá estar relacionada com as características da língua materna de cada sujeito e (4) as representações culturais adquiridas sobre a altura tonal poderão influenciar, quer a produção da fala, quer a percepção do respectivo padrão sonoro.

Para Deutsch (2008 p. 595) a percepção de altura relacionada com a dimensão circular dos sons complexos não é um fenómeno exclusivo das relações intervalares de oitava (e.g., $1f_0$, $2f_0$, $4f_0$, $8f_0$, $16f_0$, etc.). Também pode ocorrer em sequências em que os sons complexos são constituídos por componentes que formam séries harmónicas completas (e.g., $1f_0$, $2f_0$, $3f_0$, $4f_0$, $5f_0$, e assim sucessivamente). Para demonstrar a sua

asserção criou uma escala dividida em 12 partes iguais constituída por 12 sons complexos. Cada som complexo era formado pelos seis primeiros harmónicos da série harmónica completa. Ao longo da escala, à medida que a frequência fundamental dos sons desce um semi-tom, as amplitudes dos harmónicos ímpares diminuem 3.5 dB, em relação aos harmónicos pares. Criou 132 pares de sons. Cada par forma um intervalo de segunda menor. Apresentou os pares de forma aleatória. Perguntou a dezasseis sujeitos que determinassem, se o par de sons formava um padrão ascendente ou descendente, em termos de percepção de altura sonora. Nos resultados obtidos, constatou que o som com a frequência fundamental mais baixa era ouvido como se tivesse sido transportado uma oitava acima. Desta maneira, demonstrou que a percepção circular de altura também ocorre em sons constituídos por componentes que formam séries harmónicas completas, não sendo por isso, exclusiva de sons complexos cujas componentes se encontram à distância de oitava.

VII IMPLEMENTAÇÃO DA ILUSÃO SONORA DA ESCALA DE SHEPARD EM SÍNTESE DE SOM DIGITAL

Neste Capítulo, vamos implementar a Ilusão Sonora da Escala de Shepard em síntese de som digital. Referimo-nos concretamente, à sua reconstituição e reprodução áudio.

VII.1 Reconstituição da Ilusão Sonora da Escala de Shepard

Na reconstituição da Ilusão Sonora da Escala de Shepard utilizou-se um computador MacBook Intel Core 2 Duo de 2GHz, o Sistema Operativo Mac OSX (versão 10.4.11) e a aplicação *Pure Data* (versão 0.39.3-extended).

O *Pure Data* é uma linguagem de programação gráfica de acesso livre compatível com os principais Sistemas Operativos. A sua versão *extended* (*Pd-extended*) para áudio, vídeo e processamento gráfico em tempo real é semelhante ao *Max/MSP/Jitter*. O *Pure Data*, originalmente desenvolvido por Miller Puckette (1987) e pelo Ircam, é o terceiro grande ramo da família das linguagens de programação conhecidas como *Max* (*Max/FTS*, *ISPW Max*, *Max/MSP*, *jMax*, etc.).²⁰

Para determinar as frequências e as amplitudes individuais de todos os sons implementaram-se em *Pure Data* as Equações 2, 3 e 4. Na implementação das respectivas equações foram realizadas as seguintes etapas metodológicas:

- utilizou-se o objecto “*expr*” para realizar as operações existentes nas equações;
- programaram-se os *patches* de maneira a determinar os valores automaticamente e a extraí-los sob a forma de listas.

²⁰ <http://jmax.sourceforge.net/history.html>

VII.2 Determinação Automática das Frequências Individuais

A implementação da Equação 2 encontra-se representada na Figura 37.

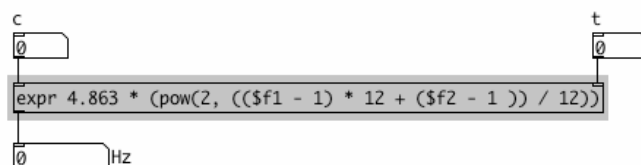


Figura 37. Implementação da Equação 2. Para $\ell=1$ introduzir na entrada do lado esquerdo do objecto “*expr*” os valores de \mathcal{C} (i.e., de um a dez); para $\ell=2$ realizar o mesmo procedimento, e assim sucessivamente até $\ell=12$.²¹

Os procedimentos representados na Figura 37, foram automatizados de maneira a extrair as frequências individuais sob a forma de listas (Figura 38 e Figura 39).

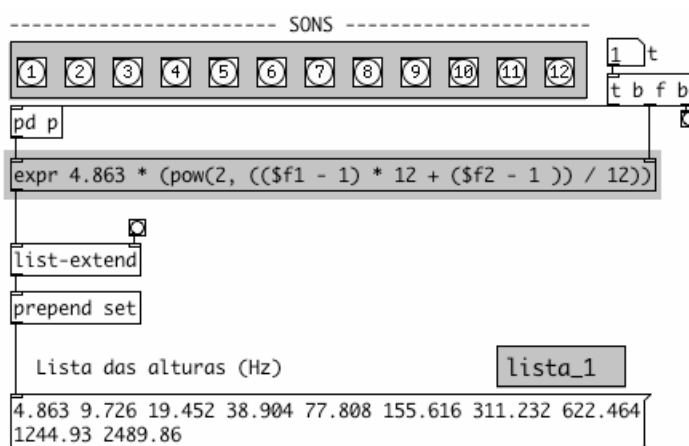


Figura 38. Determinação e extracção automática das frequências individuais, sob a forma de listas.²²

²¹ A implementação pode ser consultada no DVD em anexo, no directório *patches*, com o nome *03.equacao_2.pd*

²² O processo de extracção automática das frequências individuais, sob a forma de listas pode ser consultado no DVD em anexo, no directório *patches*, com o nome *04.extracao_automatica_frequencias_individuais.pd*

Figura 39. Listas das frequências individuais obtidas a partir do procedimento representado na Figura 38.

1	4.863 9.726 19.452 38.904 77.808 155.616 311.232 622.464 1244.93 2489.86	7	6.87732 13.7546 27.5093 55.0186 110.037 220.074 440.148 880.297 1760.59 3521.19
2	5.15217 10.3043 20.6087 41.2174 82.4347 164.869 329.739 659.478 1318.95 2637.91	8	7.28627 14.5725 29.1451 58.2901 116.58 233.161 466.321 932.642 1865.28 3730.57
3	5.45853 10.9171 21.8341 43.6683 87.3365 174.673 349.346 698.692 1397.38 2794.77	9	7.71953 15.4391 30.8781 61.7563 123.512 247.025 494.05 988.1 1976.2 3952.4
4	5.78311 11.5662 23.1325 46.2649 92.5298 185.06 370.119 740.239 1480.48 2960.95	10	8.17856 16.3571 32.7142 65.4285 130.857 261.714 523.428 1046.86 2093.71 4187.42
5	6.127 12.254 24.508 49.016 98.0319 196.064 392.128 784.256 1568.51 3137.02	11	8.66488 17.3298 34.6595 69.319 138.638 277.276 554.552 1109.1 2218.21 4436.42
6	6.49133 12.9827 25.9653 51.9306 103.861 207.722 415.445 830.89 1661.78 3323.56	12	9.18012 18.3602 36.7205 73.441 146.882 293.764 587.528 1175.06 2350.11 4700.22

Figura 39. Listas das frequências individuais obtidas a partir do procedimento representado na Figura 38.

Para uma melhor visualização dos valores das frequências individuais extraídas, elaborou-se a Tabela 2.

Tabela 2. Apresentação das frequências individuais dos Sons de Shepard, que se encontram representadas nas listas da Figura 39. As linhas (s) correspondem à variável (t) da Equação 2.

Número de Sons (s)	Número de componentes (c) e frequências individuais (Hz)									
	c1	c2	c3	c4	c5	c6	c7	c8	c9	c10
s1	4.863	9.726	19.452	38.904	77.808	155.616	311.232	622.464	1244.93	2489.86
s2	5.15217	10.3043	20.6087	41.2174	82.4347	164.869	329.739	659.478	1318.95	2637.91
s3	5.45853	10.9171	21.8341	43.6683	87.3365	174.673	349.346	698.692	1397.38	2794.77
s4	5.78311	11.5662	23.1325	46.2649	92.5298	185.06	370.119	740.239	1480.48	2960.95
s5	6.127	12.254	24.508	49.016	98.0319	196.064	392.128	784.256	1568.51	3137.02
s6	6.49133	12.9827	25.9653	51.9306	103.861	207.722	415.445	830.89	1661.78	3323.56
s7	6.87732	13.7546	27.5093	55.0186	110.037	220.074	440.148	880.297	1760.59	3521.19
s8	7.28627	14.5725	29.1451	58.2901	116.58	233.161	466.321	932.642	1865.28	3730.57
s9	7.71953	15.4391	30.8781	61.7563	123.512	247.025	494.05	988.1	1976.2	3952.4
s10	8.17856	16.3571	32.7142	65.4285	130.857	261.714	523.428	1046.86	2093.71	4187.42
s11	8.66488	17.3298	34.6595	69.319	138.638	277.276	554.552	1109.1	2218.21	4436.42
s12	9.18012	18.3602	36.7205	73.441	146.882	293.764	587.528	1175.06	2350.11	4700.22

Como as frequências individuais foram calculadas de forma logarítmica, se extraíssemos (através dos procedimentos descritos na Figura 37) as frequências a um 13.º som da Escala de Shepard, verificaríamos que teriam os mesmos valores das frequências do primeiro som da escala (s1). Constatamos assim, que na determinação das frequências das componentes nos encontramos perante um ciclo interminável. Ou seja, depois de completarmos uma oitava regressamos sempre aos valores das frequências iniciais.

Posteriormente, armazenou-se cada uma das listas num objecto chamado “*table*”. Para visualizar os nomes e as oitavas das notas musicais correspondentes às frequências individuais introduziram-se os valores de cada “*table*” num objecto chamado “*pitch*” (Figura 40).

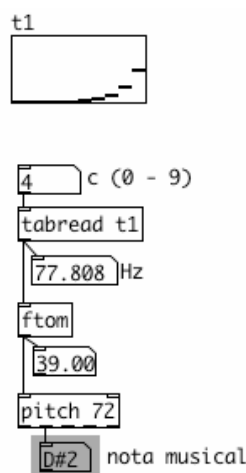


Figura 40. Processo de visualização das notas musicais correspondentes às frequências individuais que se encontram representadas na Tabela 2.²³

Aplicou-se a todos os sons da escala o procedimento representado na Figura 40 e obteve-se a Tabela 3, com os nomes e as oitavas das respectivas notas musicais.

Tabela 3. Notas musicais das frequências individuais obtidas a partir do procedimento representado na Figura 40.

Número de Sons (s)	Notas musicais das componentes (c)									
	c1	c2	c3	c4	c5	c6	c7	c8	c9	c10
s1	D#-2	D#-1	D#0	D#1	D#2	D#3	D#4	D#5	D#6	D#7
s2	E-2	E-1	E0	E1	E2	E3	E4	E5	E6	E7
s3	F-2	F-1	F0	F1	F2	F3	F4	F5	F6	F7
s4	F#-2	F#-1	F#0	F#1	F#2	F#3	F#4	F#5	F#6	F#7
s5	G-2	G-1	G0	G1	G2	G3	G4	G5	G6	G7
s6	G#-2	G#-1	G#0	G#1	G#2	G#3	G#4	G#5	G#6	G#7
s7	A-2	A-1	A0	A1	A2	A3	A4	A5	A6	A7
s8	A#-2	A#-1	A#0	A#1	A#2	A#3	A#4	A#5	A#6	A#7
s9	B-2	B-1	B0	B1	B2	B3	B4	B5	B6	B7
s10	C-1	C0	C1	C2	C3	C4	C5	C6	C7	C8
s11	C#-1	C#0	C#1	C#2	C#3	C#4	C#5	C#6	C#7	C#8
s12	D-1	D0	D1	D2	D3	D4	D5	D6	D7	D8

²³ O processo de visualização pode ser consultado no DVD em anexo, no directório *patches*, com o nome *05.notas_musicais.pd*

No final deste procedimento obteve-se a escala de D#, dividida em 12 partes iguais (escala cromática de D#).

VII.3 Determinação Automática das Amplitudes Individuais

A implementação das Equações 3 e 4 encontra-se representada na Figura 41.

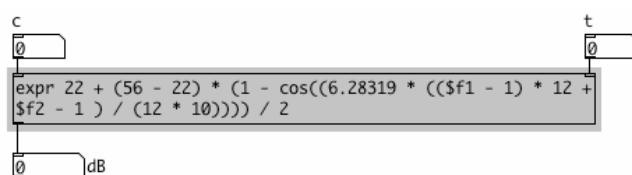


Figura 41. Implementação das Equações 3 e 4. Para $t = 1$ introduzir na entrada do lado esquerdo do objecto “*expr*” os valores de C (i.e., de um a dez); para $t = 2$ realizar o mesmo procedimento, e assim sucessivamente até $t = 12$.²⁴

Os procedimentos representados na Figura 41, foram automatizados de maneira a extrair as amplitudes individuais sob a forma de listas (Figura 42).

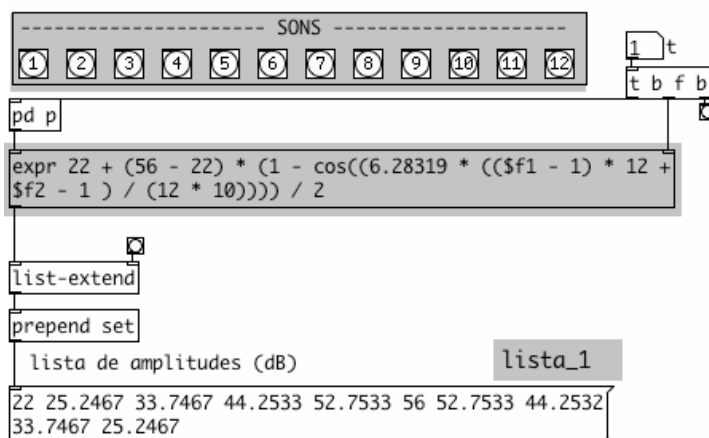


Figura 42. Determinação automática das amplitudes individuais a partir do procedimento representado na Figura 41.²⁵

²⁴ A implementação pode ser consultada no DVD em anexo, no directório *patches*, com o nome *06.equacoes_3_e_4.pd*

²⁵ A determinação automática das amplitudes individuais pode ser consultada no DVD em anexo, no directório *patches*, com o nome *07.extracao_automatica_amplitudes_individuais.pd*

No final destes procedimentos obtiveram-se 12 listas, cujos valores se encontram representados na Figura 43.

1	22 25.2467 33.7467 44.2533 52.7533 56 52.7533 44.2532 33.7467 25.2467	7	22.832 29.0077 39 48.9924 55.168 55.1679 48.9923 38.9999 29.0076 22.832
2	22.0233 25.7885 34.6001 45.0923 53.2574 55.9767 52.2114 43.3999 32.9077 24.7426	8	23.1291 29.7411 39.8897 49.6985 55.4207 54.8708 48.2588 38.1102 28.3015 22.5792
3	22.0931 26.3665 35.4655 45.9145 53.7225 55.9069 51.6334 42.5344 32.0854 24.2775	9	23.4697 30.5 40.777 50.3752 55.6285 54.5303 47.5 37.223 27.6247 22.3715
4	22.2093 26.9792 36.3406 46.7179 54.1471 55.7907 51.0208 41.6593 31.2821 23.8529	10	23.8529 31.2822 41.6594 51.0208 55.7907 54.1471 46.7178 36.3406 26.9791 22.2093
5	22.3715 27.6248 37.223 47.5 54.5303 55.6285 50.3752 40.7769 30.4999 23.4697	11	24.2776 32.0855 42.5345 51.6335 55.9069 53.7224 45.9145 35.4655 26.3665 22.0931
6	22.5793 28.3016 38.1103 48.2589 54.8709 55.4207 49.6984 39.8896 29.7411 23.1291	12	24.7426 32.9078 43.3999 52.2115 55.9767 53.2574 45.0922 34.6 25.7885 22.0233

Figura 43. Listas das amplitudes individuais obtidas a partir dos procedimentos representados na Figura 42.

Para uma melhor visualização das amplitudes individuais elaborou-se a Tabela 4.

Tabela 4. Amplitudes individuais das componentes.

Número de Sons (s)	Amplitudes individuais (dB)									
	c1	c2	c3	c4	c5	c6	c7	c8	c9	c10
s1	22	25.2467	33.7467	44.2533	52.7533	56	52.7533	44.2532	33.7467	25.2467
s2	22.0233	25.7885	34.6001	45.0923	53.2574	55.9767	52.2114	43.3999	32.9077	24.7426
s3	22.0931	26.3665	35.4655	45.9145	53.7225	55.9069	51.6334	42.5344	32.0854	24.2775
s4	22.2093	26.9792	36.3406	46.7179	54.1471	55.7907	51.0208	41.6593	31.2821	23.8529
s5	22.3715	27.6248	37.223	47.5	54.5303	55.6285	50.3752	40.7769	30.4999	23.4697
s6	22.5793	28.3016	38.1103	48.2589	54.8709	55.4207	49.6984	39.8896	29.7411	23.1291
s7	22.832	29.0077	39	48.9924	55.168	55.1679	48.9923	38.9999	29.0076	22.832
s8	23.1291	29.7411	39.8897	49.6985	55.4207	54.8708	48.2588	38.1102	28.3015	22.5792
s9	23.4697	30.5	40.777	50.3752	55.6285	54.5303	47.5	37.223	27.6247	22.3715
s10	23.8529	31.2822	41.6594	51.0208	55.7907	54.1471	46.7178	36.3406	26.9791	22.2093
s11	24.2776	32.0855	42.5345	51.6335	55.9069	53.7224	45.9145	35.4655	26.3665	22.0931
s12	24.7426	32.9078	43.3999	52.2115	55.9767	53.2574	45.0922	34.6	25.7885	22.0233

Comparando a Tabela 2 e a Tabela 4, verifica-se que à medida que as frequências das componentes aumentam, a amplitude das componentes mais baixas vai aumentando de forma gradual, desde a primeira até à sexta componente, enquanto que a amplitude das componentes mais altas vai diminuindo progressivamente, desde a sexta até à décima componente.

Para visualizar a forma da envolvente de amplitude, armazenaram-se os valores de cada lista (Figura 43) em objectos “table”.

Constatou-se, que as envolventes de amplitude apresentam a configuração de uma curva em forma de sino, ou de Curva Guassiana (Figura 44).

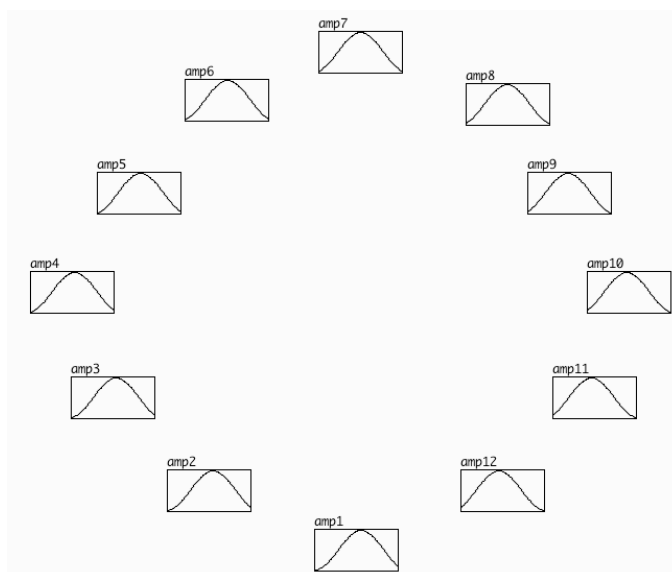


Figura 44. Representação gráfica das envolventes de amplitude dos sons que constituem a Ilusão Sonora da Escala de Shepard.

Observa-se, que as envolventes de amplitude apresentam formas similares entre si. No entanto, como foram ajustadas de forma ponderada a cada um dos sons que constituem a ilusão, as suas configurações vão sofrendo ligeiras alterações. Compare-se, por exemplo, a envolvente “*amp1*” com a envolvente “*amp12*” (Figura 45).

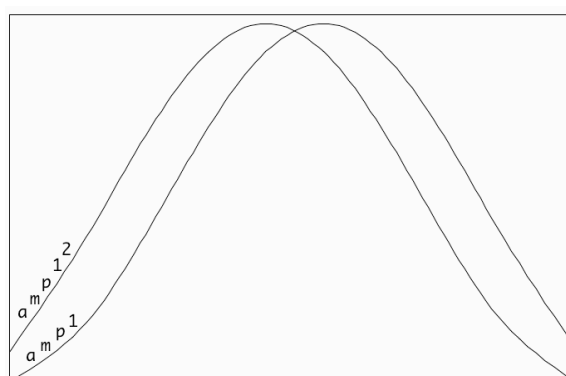


Figura 45. Comparação entre as envolventes de amplitude “*amp1*” e “*amp12*”. As envolventes encontram-se sobrepostas.

Através da Figura 45, constata-se uma ligeira diferença, quer no declive ascendente, quer no declive descendente da curva das respectivas envolventes. No

entanto, se colocássemos a envolvente “*ampl*” em frente a um espelho plano verificaríamos, que a sua imagem reflectida corresponderia à imagem da envolvente “*ampl2*”. O que significa, que as envolventes são simétricas. Ou seja, se invertêssemos uma delas e a sobrepusemos à outra constataríamos que se ajustavam e/ou seriam idênticas. Como as envolventes foram calculadas de forma logarítmica, se extraíssemos (através dos procedimentos descritos na Figura 41) a envolvente espectral a um 13.º som da Escala de Shepard, verificaríamos que esta, seria igual à envolvente “*ampl*”. Constatamos assim, que na determinação das envolventes de amplitude nos encontramos também perante um ciclo interminável. Depois de completarmos uma volta regressamos sempre aos valores de amplitude iniciais.

VII.4 Reprodução Áudio da Ilusão Sonora da Escala de Shepard

A reprodução áudio da Ilusão da Sonora da Escala de Shepard equivale a um processo de síntese aditiva.

Para as frequências individuais foram utilizados dez osciladores que se encontram representados pelo objecto “*osc~*” (i.e., um oscilador para cada componente de cada som complexo).

No que respeita às amplitudes individuais foram utilizados dez operadores de multiplicação áudio que se encontram representados pelos objectos “**~*” (Figura 46).

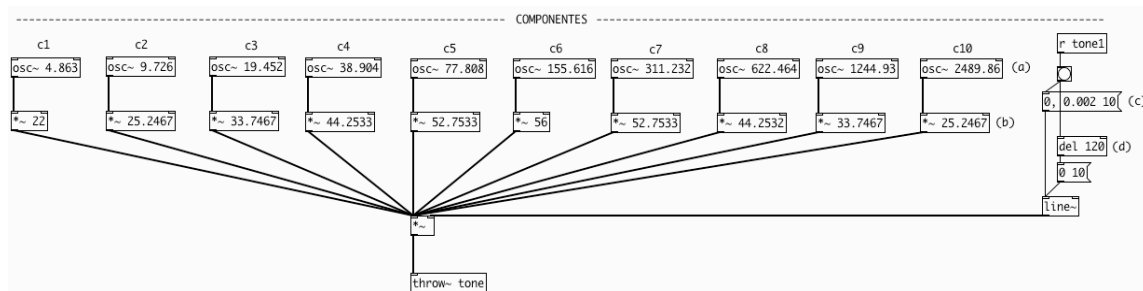


Figura 46. Síntese do primeiro som da Ilusão Sonora da Escala de Shepard.²⁶ O ponto (a) representa as frequências individuais (Hz) das componentes e o ponto (b) as respectivas amplitudes (dB). Por sua vez, os pontos (c) e (d) representam a envolvente de amplitude geral aplicada ao som. A amplitude geral do som consiste no seguinte: a amplitude sobe linearmente de zero até 0.002 em 10 milissegundos; após 120 milissegundos, desce de 0.002 até zero em 10 milissegundos.

²⁶ O exemplo pode ser consultado no DVD em anexo, no directório *patches*, com o nome *08.sintese_som_shepard.pd*

Aplicou-se o processo de síntese aditiva representado na Figura 46, a todos os sons complexos que constituem a escala. Atribuí-se a cada um deles as frequências (Tabela 2) e as amplitudes (Tabela 4) correspondentes.

Após este procedimento, criou-se uma aplicação com uma interface de utilizador para reproduzir auditivamente a ilusão. A interface de utilizador encontra-se representada na Figura 47.

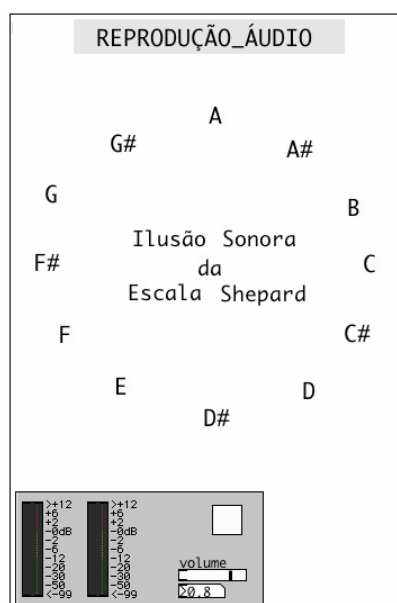


Figura 47. Interface de utilizador da reprodução áudio da Ilusão Sonora da Escala de Shepard.²⁷

VII.5 Percepção Auditiva da Reprodução Áudio da Ilusão Sonora da Escala de Shepard

Apesar dos resultados auditivos obtidos a partir da reprodução áudio da Ilusão Sonora da Escala de Shepard terem sido satisfatórios constatámos no entanto, que não produziram plenamente o efeito da ilusão que esperávamos obter, como aconteceu por exemplo, quando escutámos pela primeira vez o ficheiro áudio da gravação original de Shepard.

²⁷ A interface de utilizador para reproduzir a Ilusão Sonora da Escala de Shepard, pode ser consultada no DVD em anexo, no directório *patches*, com o nome *09_reproducao_audio.pd*

Na reprodução é audível um pequeno salto descendente no momento em que a escala regressa à tônica. Para além disso, há a proeminência indesejada de uma componente no registo sonoro agudo que vai surgindo progressivamente, sobretudo a partir do meio da escala (i.e., do sétimo som ou da nota musical A) e que acaba por se extinguir sempre que a escala regressa à tônica.

Esta constatação é semelhante à percepção auditiva que os ouvintes com treino musical revelaram na demonstração de Shepard (1964, p. 2349).

Para nos certificarmos que não ocorreu nenhum erro durante o processo de reconstituição e reprodução áudio da ilusão realizámos uma análise espectral ao Som de Shepard.

VIII ANÁLISE ESPECTRAL DO SOM DE SHEPARD

Para obter uma representação gráfica das envolventes de amplitude, no domínio das frequências realizou-se uma análise espectral a todos os sons da escala, recorrendo à análise de Fourier.

VIII.1 Análise de Fourier

A análise de Fourier é um conjunto de técnicas matemáticas inventadas pelo Francês Jean Baptiste Joseph Fourier (1768-1830). Estas técnicas baseiam-se na decomposição de sinais em ondas sinusoidais (Smith 1999, p. 141).

A análise de Fourier permite converter um sinal no domínio do tempo na sua representação equivalente no domínio das frequências (Kester 2000, p. 5.1).

A decomposição de um sinal apresenta a vantagem de permitir trabalhar individualmente com cada um dos elementos que o constituem, em vez de se lidar com o sinal inteiro.

Os sinais podem ser contínuos ou discretos, periódicos ou aperiódicos. A combinação destas características originam quatro categorias de análise de Fourier (e.g., Transformada de Fourier, séries de Fourier, transformada de Fourier em tempo discreto e a transformada discreta de Fourier) (Kester 2000, p. 5.1). Destas categorias, destaca-se a transformada discreta de Fourier (DFT), pelo facto, de se poder aplicar a sinais discretos e periódicos. Assim, como os computadores processam informação finita, periódica e discreta, a DFT é a única análise que se pode utilizar no Processamento de Som Digital (*Digital Sound Processing*, DSP) (Smith 1999, p. 145). Para além disso, como o sinal é periódico permite ser decomposto numa soma de ondas sinusoidais (Kester 2000, p. 5.3).

VIII.2 Breve Descrição de uma DFT

A DFT pode ser subdividida em duas versões: a versão real e a versão complexa. A versão real torna-se na prática mais simples, porque utiliza somente números reais, enquanto que a versão complexa, por usar números complexos requer maior número de cálculos e conseqüentemente torna-se menos eficaz em termos de processamento (Smith 1999, p. 146).

Considera-se, que o sinal de entrada se encontra no domínio do tempo e que o sinal de saída, se encontra no domínio das frequências. O domínio das frequências contém exactamente a mesma informação existente no domínio do tempo, só que em formato diferente (Smith 1999, p. 147). Basta conhecer um dos domínios para calcular o outro.

Se tivermos um sinal no domínio do tempo, o processo de calcular o seu equivalente no domínio das frequências chama-se decomposição, análise ou simplesmente DFT. Se por outro lado, tivermos um sinal no domínio das frequências, o processo de cálculo do seu equivalente no domínio do tempo chama-se síntese ou DFT inversa (IDFT). Ambos os processos (análise e síntese) podem ser representados por equações e algoritmos informáticos (Smith 1999, p. 147).

A DFT transforma N amostras do sinal de entrada em duas vezes $N/2 + 1$ amostras de sinal de saída. O sinal de entrada contém a informação que vai ser analisada, enquanto que o sinal de saída contém as amplitudes das ondas sinusoidais (Smith 1999, p. 146).

Pode-se atribuir qualquer número inteiro positivo a N , no entanto, é comum atribuir-lhe valores representados por potências de base dois (2^n). Este pormenor, deve-se ao facto, de o algoritmo mais comum para calcular a DFT, a transformada rápida de Fourier (FFT) operar com valores de N representados por potências de base dois. Por exemplo: 2^2 , 2^3 , 2^4 , 2^5 , 2^6 , e assim sucessivamente (Smith 1999, p. 147).

VIII.3 Transformada rápida de Fourier (FFT)

Apesar de Cooley e Tukey (1965) terem introduzido no âmbito da informática digital o algoritmo da FFT, através do Artigo “*An Algorithm for the machine calculation of complex Fourier Series*”, *Mathematics Computation*, Vol. 19, 1965, pp. 297-301, já o matemático alemão Karl Friedrich Gauss (1777-1855) o usava, pelo menos, há um século atrás (Smith 1999, p. 225).

O trabalho de Gauss não foi muito divulgado, porque nessa altura ainda não se tinha inventado o computador digital, e para além disso, efectuar os cálculos manualmente tornava-se numa longa e árdua tarefa.

Cooley e Tukey têm pelo menos o mérito de terem divulgado a FFT na altura certa, ou seja, no início da revolução digital (Smith, 1999, p. 225).

A FFT decompõe um sinal do domínio do tempo de N amostras em N sinais, cada um deles constituído por uma única amostra (Smith 1999, p. 228). Na FFT são necessárias $\log_2 N$ etapas para realizar a análise. Por exemplo, uma FFT de um sinal de $N = 8$ amostras (2^3 , i.e., 3 Bits) requer três etapas, um sinal de 512 amostras (2^9 , i.e., 9 Bits) requer nove etapas, um sinal de 4096 amostras (2^{12} , i.e., 12 Bits) requer doze etapas, e assim sucessivamente (Smith 1999, p. 229).

A FFT, pode ser modificada para calcular a versão real da DFT. Para tal, utiliza-se outro algoritmo, que se chama a transformada rápida de Fourier real (RFFT). Este novo algoritmo é 30% mais rápido do que o algoritmo da FFT (Smith 1999, p. 239). O seu cálculo requer somente metade do tempo e de processamento de uma FFT (Smith 1999, p. 241). Através de uma RFFT obtém-se uma FFT de $N/2$ amostras.

Sai do âmbito desta dissertação, fazer uma descrição pormenorizada de uma RFFT. Por essa razão, limitamo-nos a descrever, de forma sucinta os principais passos que efectuámos na análise ao som de Shepard, através de uma RFFT realizada em *Pure Data*.²⁸

VIII.4 Análise do Som de Shepard através de uma transformada rápida de Fourier real (RFFT) realizada em Pure Data

Começámos por determinar a resolução da análise. O valor da frequência mais baixa do Som de Shepard, ou seja, a frequência da primeira componente do primeiro som da Ilusão Sonora da Escala de Shepard é igual a 4.863 Hz (Tabela 2). Com uma taxa de amostragem (*sampling rate*) de 44100 Hz, para se alcançar uma resolução, que permita analisar essa frequência deve-se utilizar um bloco de análise (*blocksize*) com o tamanho de 16384 amostras. Assim, dividindo a taxa de amostragem pelo bloco de análise (SR/BS) obtém-se a resolução da RFFT. Neste caso, terá o valor de 2.6916504 Hz.

²⁸ Para uma descrição pormenorizada de uma RFFT em *Pure Data* consultar os seguintes autores: Barkenecht (2007), Kreidler (2007) e Puckette (2007).

Depois, para evitarmos um processo de fuga espectral (*leakage*), que ocorre nas frequências que não estabelecem relações harmónicas inteiras com a frequência fundamental, multiplicámos os valores do sinal de entrada por uma lista de números positivos. Para isso, utilizámos uma janela *Hanning* (i.e., uma função matemática) (Kester 2000, p. 5.22). Este processo tem o nome de *Windowing*. Para determinar a respectiva janela recorreu-se à Equação 5 (Puckette 2007, p. 271; Kester 2000, p. 5.23).

$$W(n) = 0.5 - 0.5 \cos\left(\frac{2\pi n}{N}\right), \quad (5)$$

Onde:

- n representa o eixo horizontal; o seu valor está compreendido entre 0 e $N-1$;
- N representa o tamanho do *BS*.

A Figura 48, exemplifica a implementação áudio da Equação 5.

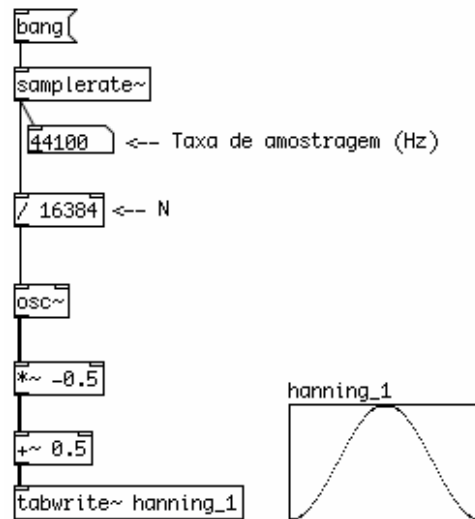


Figura 48. Implementação áudio da Equação 5 e visualização da janela *Hanning*.²⁹

²⁹ A implementação áudio da Equação 5 pode ser consultada no DVD em anexo, no directório *patches*, com o nome *10.implementacao_audio_equacao_5.pd*

A multiplicação do sinal de entrada pelos valores da janela *Hanning* permite que o sinal nas suas extremidades seja sempre igual a zero fazendo com que a análise seja sempre constante e periódica (Kester 2000, p. 5.22).

Como complemento do processo *Windowing*, a janela *Hanning* precisa de ser sobreposta ao longo do tempo. Para isso, basta atribuir ao objecto “*block~*” (através de um segundo argumento) o número de janelas que pretendemos utilizar por segundo. Neste caso, com um objecto “*block~ 16384 4*”, por cada bloco de análise de 16384 amostras a janela é lida quatro vezes por intervalos regulares de 4096 amostras (Puckette 2007, p. 287).

O *Pure Data* possui um objecto áudio específico para realizar a RFFT. O objecto chama-se “*rfft~*”. Nesse objecto, a parte imaginária do sinal de entrada é reduzida a zero. Só a saída (*outlet*) do lado esquerdo do objecto é que se encontra numericamente preenchido. Os valores atribuídos à saída do lado direito são determinados por simetria (Puckette 2007, p. 287). Com este procedimento, reduz-se o tempo de processamento.

O objecto “*rfft~*” envia através das suas saídas a parte real e a parte imaginária do sinal analisado, de acordo, com a seguinte configuração (Barkenecht 2007, pp. 1-2):

- saída esquerda: parte real → cosenos → fases;
- saída direita: parte imaginária → senos → amplitudes.

Basicamente, o objecto “*rfft~*” retorna as frequências, as fases e as amplitudes de um conjunto de sinusoidais. Desta maneira, os osciladores ao serem adicionados permitem representar o sinal original no domínio das frequências.

A RFFT gera informação de dados da seguinte maneira: $BS/2$ osciladores. Assim, com um BS de 16384 Hz obtêm-se as frequências, as fases e as amplitudes de 8192 osciladores. O último oscilador, deverá ter a frequência final igual a $(BS/2) * SR/BS$, em que $SR/2$ corresponde à frequência de *Nyquist*. Valores acima da frequência de *Nyquist* poderão originar um fenómeno conhecido em DSP por Distorção ou *Foldover* (Puckette 2007, p. 60).

Os valores das fases e das amplitudes na RFFT aparecem sob a forma de valores de senos e cosenos. Os valores emitidos pelos *outlets* do objecto “*rfft~*” constituem pares de amostras (números). Os pares de amostras extraídos são convertidos nas amplitudes através da Equação 6 (Barkenecht 2007, p. 2).

$$\sqrt{(re^2 + im^2)}, \quad (6)$$

Onde:

- *re* representa a parte real;
- *im* representa a parte imaginária.

Desta maneira, passa-se de uma representação cartesiana ou rectangular para uma representação trigonométrica ou polar (Barkenecht 2007, p. 2). A implementação áudio da Equação 6 está representada na Figura 49.

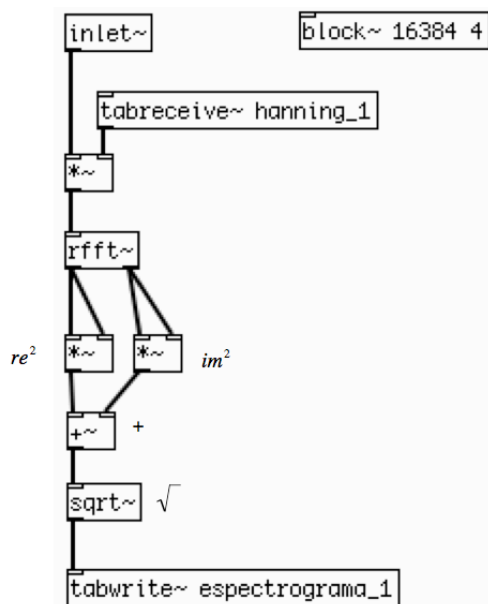


Figura 49. Implementação áudio da Equação 6.³⁰

O próximo passo consistiu na obtenção de uma representação da RFFT, sob a forma de uma tabela (i.e., espectrograma) (Kreidler 2007, p. 2). No espectrograma, a cada um dos pontos temporais corresponde o resultado de uma RFFT (Figura 50).

³⁰ A implementação áudio da Equação 6 pode ser consultada no DVD em anexo, no directório *patches*, com o nome de *11_implementacao_audio_equacao_6.pd*

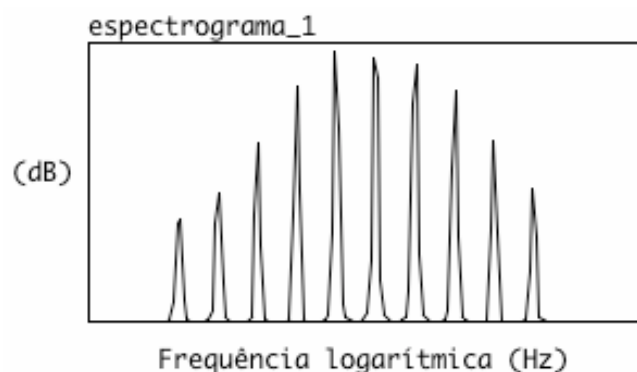


Figura 50. Espectrograma do primeiro do Som de Shepard.

Depois de efectuarmos uma RFFT a todos os Sons de Shepard verificámos que, quer as componentes, quer as respectivas envolventes espectrais se encontravam presentes e não evidenciavam qualquer tipo de erro em termos de DSP.

O facto, de termos percebido, na reprodução áudio da ilusão Sonora da Escala de Shepard, um salto descendente entre a última nota e a tónica da escala, e ainda a proeminência indesejada de uma componente no registo sonoro agudo que começa a destacar-se a partir do meio da escala e que acaba por se extinguir no final da mesma, gerou-nos uma insatisfação sonora ou musical. Para além disso, quer o Som de Shepard original, quer o som por nós reproduzido apresentavam timbres baixos, escuros, pesados e pouco brilhantes, sendo por isso considerados sonora e musicalmente pouco interessantes para serem utilizados nos nossos projectos musicais. Assim, decidimos criar a nossa própria fonte sonora. A fonte sonora proposta e as suas aplicações auditivo-musicais irão ser apresentadas ao longo da terceira parte da dissertação.

**PARTE III –
DESENVOLVIMENTO,
APLICAÇÕES E COMPOSIÇÃO**

IX CHROMASOM: UMA FONTE SONORA PARADOXAL

Nesta secção da dissertação concebemos um método original de criar uma fonte sonora digital baseada no Som de Shepard. A fonte sonora proposta é paradoxal, no sentido, em que é um som ambíguo em termos de percepção de altura sonora. Provavelmente, conseguiremos identificar a nota musical do som, mas teremos muita dificuldade em determinar a oitava a que pertence. Foi denominada Chromasom, pelo facto, de preservar a altura tonal, ou seja, a posição do som dentro da oitava.³¹ O seu nome deriva também das palavras *Tone chroma*, que constituem o termo anglo-saxónico correspondente à altura tonal.

Vamos aplicar ao Chromasom uma envolvente espectral alternativa, tanto quanto sabemos, nunca antes foi usada para reproduzir a Ilusão Sonora da Escala de Shepard, mesmo quando a ilusão em questão é aplicada a uma melodia constituída por uma sequência de saltos intervalares não regulares.

Na construção do Chromassom iremos realizar as seguintes tarefas:

- aplicação de envolventes espectrais utilizadas por outros autores;
- concepção de uma envolvente espectral alternativa que permita reproduzir a Ilusão Sonora da Escala de Shepard;
- aplicação do Chromasom a escalas musicais divididas em 24 e 48 partes iguais (i.e., escalas de 1/4 e 1/8 de tom), a acordes musicais, a progressões melódicas e harmónicas e a uma melodia constituída por uma sequência de saltos intervalares não regulares, na tentativa de verificar, se o efeito da ilusão sonora se mantém, ou seja, se a sensação de perpetuidade sonora se mantém durante a repetição dos exemplos mencionados;

³¹ Por termos alterado recentemente o nome da fonte sonora de *Our Shepard Sound* (OSS) para Chromasom é possível encontrar a designação inicial em documentos e registos multimédia publicados antes de se ter efectuado a respectiva alteração.

- utilização do Chromasom em técnicas de síntese de som digital (e.g., síntese por modulação de frequência), com o objectivo de obter timbres sonoros interessantes, para usar em composição musical.

Nas próximas secções iremos descrever o processo de determinação automática das frequências individuais, seguido do processo de determinação das envolventes espectrais usadas por outros autores. As envolventes espectrais dos outros autores vão ser aplicadas ao Chromasom, e posteriormente, testadas na reprodução da Ilusão Sonora da Escala de Shepard. Por último, apresentaremos uma aplicação com uma interface de utilizador, que permitirá escutar o resultado sonoro obtido nos processos previamente mencionados.

IX.1 Determinação das Frequências Individuais do Chromasom

Utilizou-se a Equação 2 para determinar as frequências individuais do Chromasom. No entanto, atribuí-se à constante f_{\min} da respectiva Equação o valor de 32.72 Hz. Este valor correspondente à nota musical C1. Desta maneira, deslocámos para o lado direito, no eixo x (o eixo das frequências das componentes) todo o espectro sonoro com o propósito de tornar o som mais agudo, timbricamente mais brilhante e ainda, sonora e musicalmente mais interessante, para usar em composição musical, do que o Som de Shepard original ou o Som de Shepard anteriormente reproduzido em *Pure Data*.

Após realizarmos os procedimentos de cálculo representados na Figura 38, Figura 39 e Figura 40 obtivemos a escala musical de C dividida em 12 partes iguais (escala cromática de C) (Tabela 5 e Tabela 6).

Tabela 5. Frequências individuais do Chromasom.

Número de Sons (s)	Número de componentes e frequências individuais (Hz) do Chromasom								
	c1	c2	c3	c4	c5	c6	c7	c8	c9
s1	32.72	65.44	130.88	261.76	523.52	1047.04	2094.08	4188.16	8376.32
s2	34.6656	69.3313	138.663	277.325	554.65	1109.3	2218.6	4437.2	8874.4
s3	36.727	73.4539	146.908	293.816	587.631	1175.26	2350.53	4701.05	9402.1
s4	38.9109	77.8217	155.643	311.287	622.574	1245.15	2490.29	4980.59	9961.18
s5	41.2246	82.4492	164.898	329.797	659.594	1319.19	2638.38	5276.75	10553.5
s6	43.676	87.3519	174.704	349.408	698.815	1397.63	2795.26	5590.52	11181
s7	46.2731	92.5461	185.092	370.185	740.369	1480.74	2961.48	5922.95	11845.9
s8	49.0246	98.0492	196.098	392.197	784.394	1568.79	3137.58	6275.15	12550.3
s9	51.9398	103.88	207.759	415.518	831.036	1662.07	3324.14	6648.29	13296.6
s10	55.0283	110.057	220.113	440.226	880.452	1760.90	3521.81	7043.62	14087.2
s11	58.3004	116.601	233.202	466.403	932.807	1865.61	3731.23	7462.45	14924.9
s12	61.7671	123.534	247.069	494.137	988.274	1976.55	3953.10	7906.19	15812.4

Tabela 6. Notas musicais das componentes do Chromasom.

Número de Sons (s)	Notas musicais das componentes (c) do Chromasom								
	c1	c2	c3	c4	c5	c6	c7	c8	c9
s1	C1	C2	C3	C4	C5	C6	C7	C8	C9
s2	C#1	C#2	C#3	C#4	C#5	C#6	C#7	C#8	C#9
s3	D1	D2	D3	D4	D5	D6	D7	D8	D9
s4	D#1	D#2	D#3	D#4	D#5	D#6	D#7	D#8	D#9
s5	E1	E2	E3	E4	E5	E6	E7	E8	E9
s6	F1	F2	F3	F4	F5	F6	F7	F8	F9
s7	F#1	F#2	F#3	F#4	F#5	F#6	F#7	F#8	F#9
s8	G1	G2	G3	G4	G5	G6	G7	G8	G9
s9	G#1	G#2	G#3	G#4	G#5	G#6	G#7	G#8	G#9
s10	A1	A2	A3	A4	A5	A6	A7	A8	A9
s11	A#1	A#2	A#3	A#4	A#5	A#6	A#7	A#8	A#9
s12	B1	B2	B3	B4	B5	B6	B7	B8	B9

O Chromasom é constituído por nove componentes separadas por intervalos de oitava. Usamos somente nove componentes, porque as frequências das nonas componentes, a partir do quarto som complexo da escala (D#1) ultrapassam o limite máximo relativo da audição humana de uma pessoa jovem (e.g., de 20 anos de idade), limite esse, que se situa por volta dos 20 kHz (Benson 2008, pp. 8-9; Loy 2006, p. 153).

De seguida, vamos testar na construção do Chromasom algumas das envolventes espectrais utilizadas por outros autores, na tentativa de verificar se reproduzem a Ilusão Sonora da Escala de Shepard.

IX.2 Envoltentes Espectrais Utilizadas por Outros Autores Aplicadas ao Chromasom

Aplicámos ao Chromasom as seguintes envoltentes espectrais:

IX.2.1 Envoltente Espectral de Pollack

Pollack (1978, p. 202) usou uma envoltente espectral triangular com amplitudes fixas determinadas da seguinte maneira: as amplitudes individuais são sempre as mesmas para as componentes centrais (cinco e seis), enquanto que as amplitudes das componentes, que se situam quer abaixo quer acima das frequências das componentes centrais diminuem logarithmicamente em ambas as direcções do espectro (Figura 51).

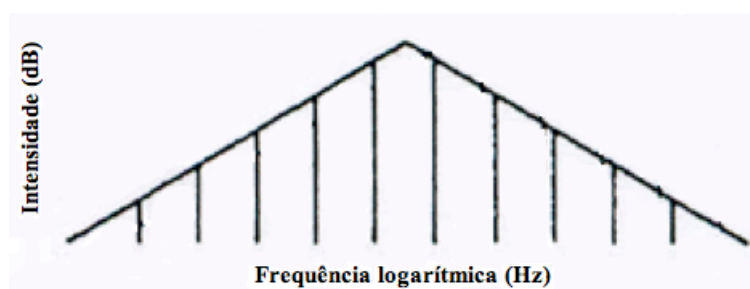


Figura 51. Envoltente espectral de Pollack. As linhas verticais representam as dez componentes que se encontram separadas por intervalos de oitava e a linha oblíqua a envoltente espectral. Adap. de Pollack (1978, p. 202).

Para determinar a envoltente espectral de Pollack construímos uma escala logarithmica a partir da Equação 7.

$$y = \log_{10}(x), \quad (7)$$

Onde:

$$y = \log x / \log 10$$

Para se obterem os valores de y (que vão representar as amplitudes) implementou-se em *Pure Data* a Equação 7 (Figura 52).

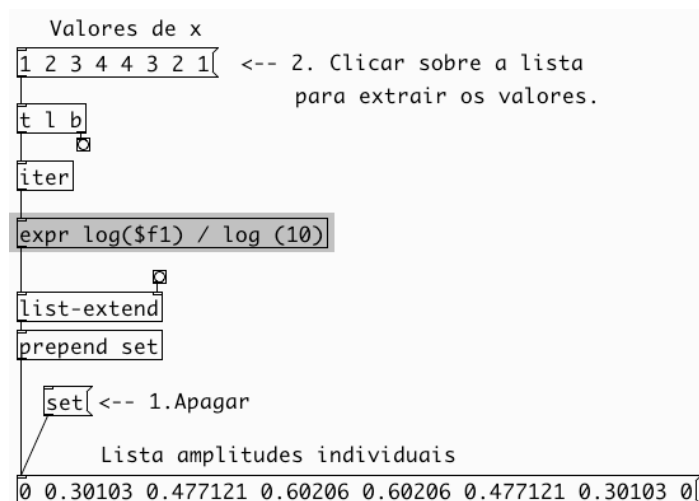


Figura 52. Implementação da Equação 7 e extracção automática das amplitudes individuais sob a forma de lista.³²

A representação gráfica da envolvente espectral obtida pode ser visualizada através da Figura 53.

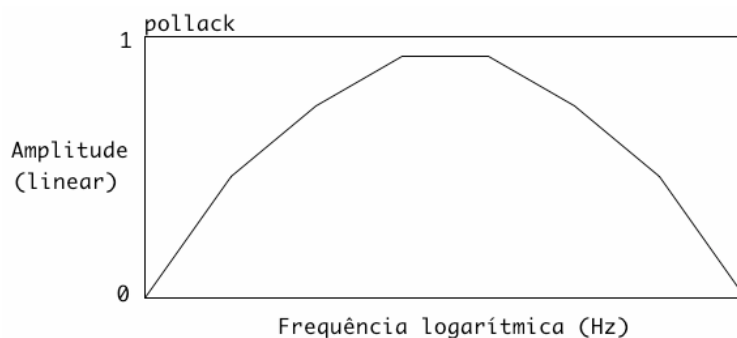


Figura 53. Representação gráfica da reprodução da envolvente espectral de Pollack.

IX.2.2 Envolvente Espectral de Burns

Burns (1981, p. 468) apresenta uma envolvente espectral triangular muito semelhante à envolvente espectral de Pollack, na qual as amplitudes individuais foram ponderadas da seguinte maneira: menos 6 dB por oitava, desde a componente central em ambas as direcções do espectro (Figura 54).

³² A implementação da Equação 7 e a representação gráfica da envolvente espectral de Pollack podem ser consultadas no DVD em anexo, no directório *patches*, com o nome *12.equacao_7_env_pollack.pd*

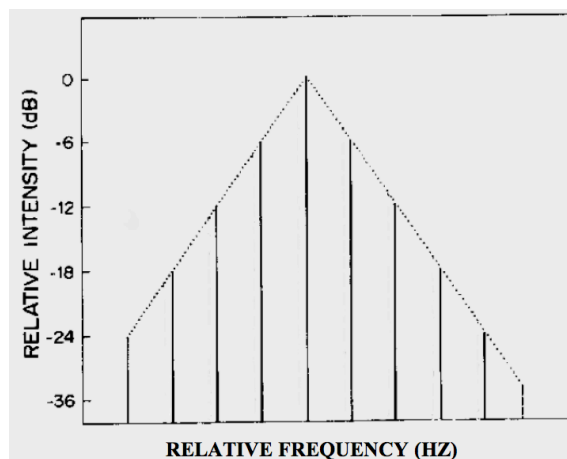


Figura 54. Envolvente espectral de Burns. No eixo x , as frequências estão representadas de forma logarítmica. As linhas verticais simbolizam as dez componentes do som e a linha constituída por pontos a envolvente espectral. Adap. de Burns (1981, p. 468).

Para reproduzir a envolvente espectral de Burns converteram-se os valores negativos das intensidades relativas das componentes, que se encontram representadas no eixo y da Figura 54, em valores positivos através do objecto “abs”. Este objecto permite obter o valor absoluto de um número. De seguida, normalizaram-se os valores extraídos numa escala linear compreendida entre zero e um, através da Equação 8.

$$f1 * a/b, \quad (8)$$

Onde:

- $f1$ representa a variável;
- a representa o limite máximo da escala normalizada;
- b representa o valor máximo positivo do conjunto de números que se pretendem normalizar.

Ambos os processos de conversão encontram-se representados na Figura 55.

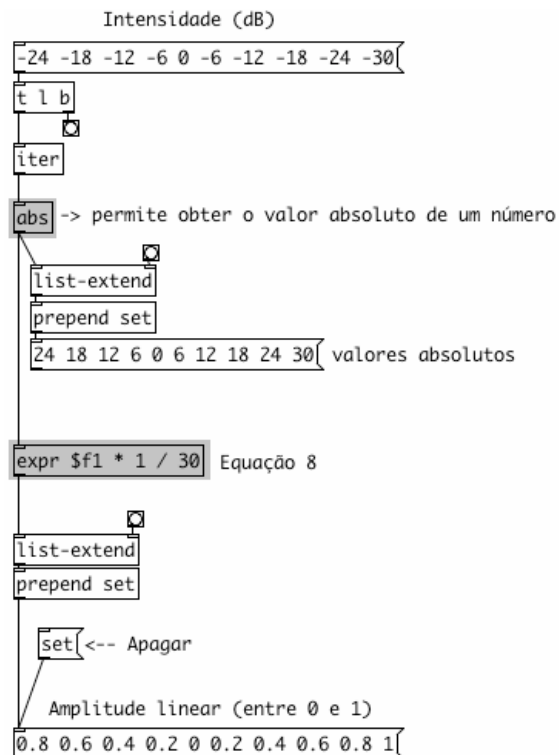


Figura 55. Conversão dos valores negativos de intensidade em positivos através do objecto “abs”. Normalização dos valores obtidos, através da Equação 8, numa escala compreendida entre zero e um.³³

A envolvente espectral obtida está representada na Figura 56.

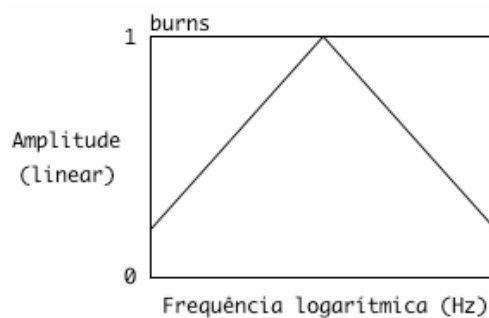


Figura 56. Representação gráfica da reprodução da envolvente espectral de Burns.

³³ O processo de conversão e a representação gráfica da envolvente espectral de Burns podem ser consultados no DVD em anexo, no directório *patches*, com o nome *I3.conversao_env_burns.pd*

IX.2.3 Envolvente Espectral de Nakajima et al.

A envolvente espectral de Nakajima *et al.*, (1988, p. 7) caracteriza-se pelo seguinte: as amplitudes das componentes centrais apresentam um valor constante, enquanto que nas componentes das extremidades o nível de pressão sonora diminui com uma pendente de -20 dB por oitava (Figura 57).

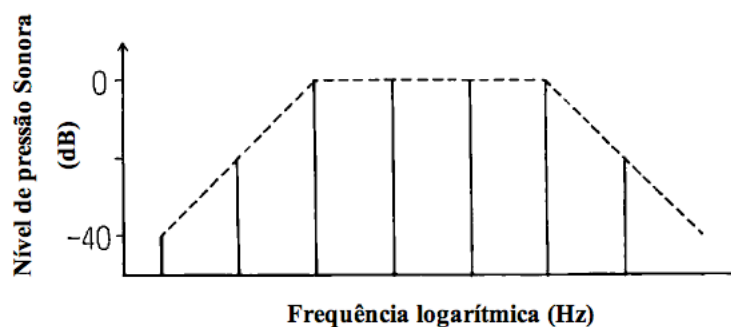


Figura 57. Envolvente espectral de Nakajima *et al.* As linhas verticais representam as componentes do som e o tracejado a envolvente espectral. Adap. de Nakajima *et al.*, (1988, p. 7).

Para reproduzir a envolvente espectral de Nakajima *et al.*, utilizaram-se os mesmos procedimentos realizados na reprodução da envolvente espectral de Burns. Primeiro, converteram-se os valores negativos das intensidades das componentes, que se encontram representadas no eixo y em valores positivos através do objecto “*abs*”. De seguida, através da Equação 8, normalizaram-se os valores extraídos numa escala linear compreendida entre zero e um. Os processos de conversão encontram-se representados na Figura 58.

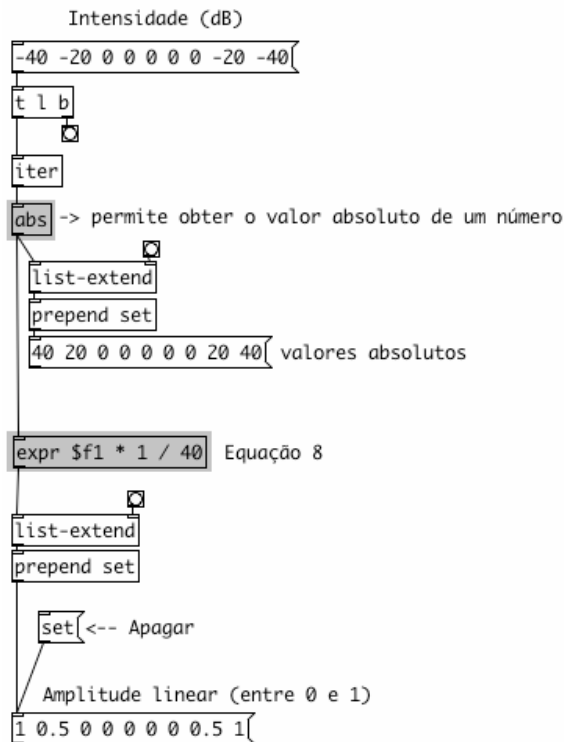


Figura 58. Conversão dos valores negativos de intensidade em positivos através do objecto “abs”. Normalização dos valores obtidos, através da Equação 8, numa escala compreendida entre zero e um.³⁴

A envolvente espectral obtida está representada na Figura 59.

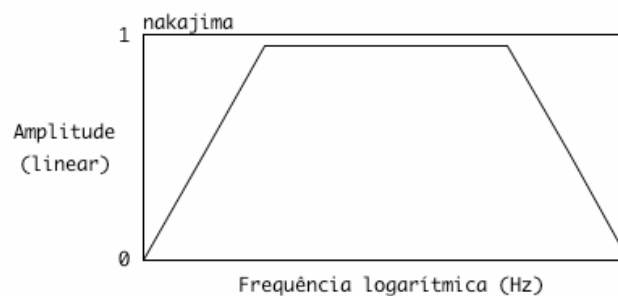


Figura 59. Representação gráfica da reprodução da envolvente espectral de Nakajima *et al.*

³⁴ O processo de conversão e a representação gráfica da envolvente espectral de Nakajima *et al.*, podem ser consultados no DVD em anexo, no directório *patches*, com o nome *14.conversao_env_nakajima.pd*

IX.2.4 Envolvente Espectral de Deutsch

A envolvente espectral de Deutsch (1986, p. 1348) encontra-se representada na Figura 60.

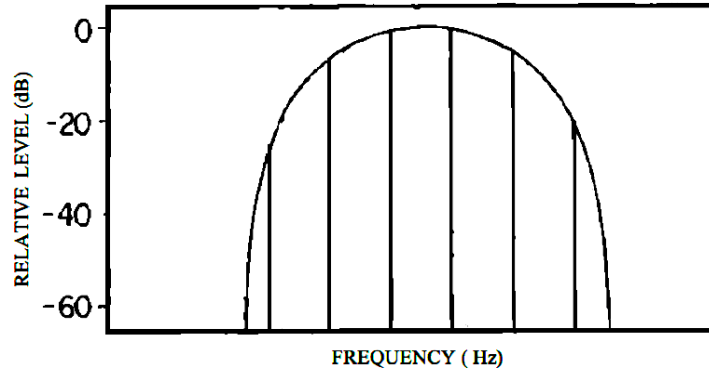


Figura 60. Envolvente espectral de Deutsch. No eixo x , as frequências estão representadas de forma logarítmica. As linhas verticais representam as seis componentes separadas por intervalos de oitava que constituem o som e a linha curva a envolvente espectral. Adap. de Deutsch et al. (1986, p. 1348).

Deutsch (1986, pp. 1347-1348; 1992, p. 392) construiu a sua envolvente espectral a partir da Equação 9.

$$A(f) = 0.5 - 0.5 \cos\left[\frac{2\pi}{y} \log_{\beta}\left(\frac{f}{f_{\min}}\right)\right], \quad (9)$$
$$f_{\min} \leq f \leq \beta^y f_{\min},$$

Onde:

- $A(f)$ é a amplitude relativa de uma componente para uma determinada frequência f Hz;
- $f_{\min} = 16.36$ Hz, representa o valor da frequência mínima em que o valor da amplitude é diferente de zero;
- β é a proporção da frequência entre as componentes sinusoidais adjacentes (e.g., para intervalos de oitava $\beta = 2$);
- $y = 9$, representa o número de componentes de cada som.

A envolvente espectral para o primeiro som abrange um total de nove oitavas. As suas frequências estão compreendidas entre f_{\min} e $512f_{\min}$.

O processo de implementação da Equação 9 e a extração automática das amplitudes individuais encontra-se representado na Figura 61.

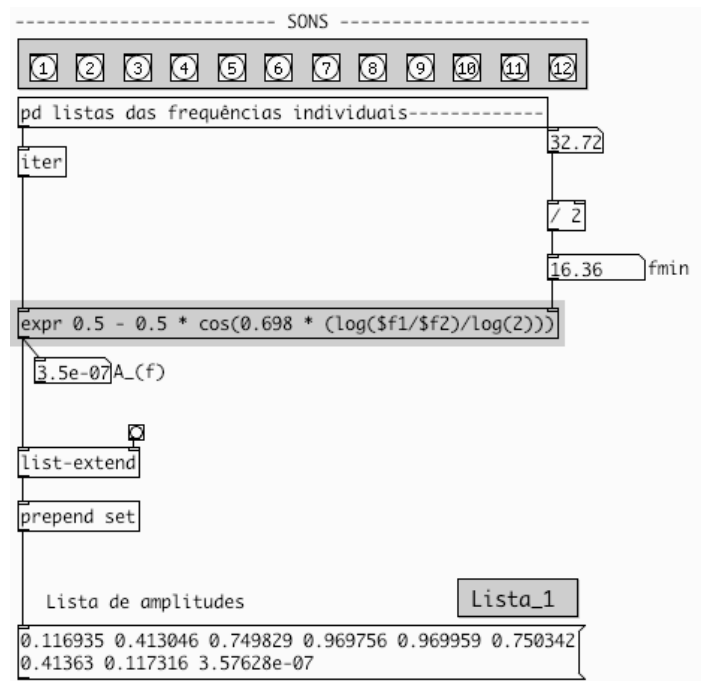


Figura 61. Implementação da Equação 9 e extração automática das listas das amplitudes individuais.³⁵

A representação gráfica da envolvente espectral de Deutsch obtida encontra-se representada na Figura 62.

³⁵ A implementação da Equação 9 e a representação gráfica da envolvente espectral de Deutsch podem ser consultadas no DVD em anexo, no directório *patches*, com o nome *15.equacao_9_env_deutsch.pd*

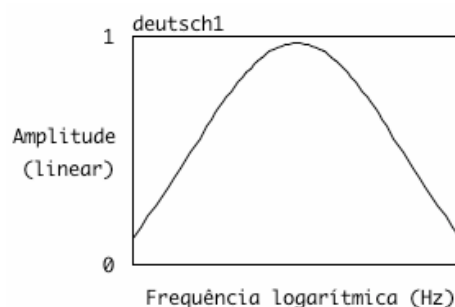


Figura 62. Representação gráfica da reprodução da envolvente espectral de Deutsch.

IX.3 Percepção Auditiva das Envolventes Espectrais

É importante realçar, que em todas as envolventes espectrais apresentadas até agora, as componentes das extremidades possuem sempre as amplitudes mais baixas, enquanto que as componentes centrais possuem sempre as amplitudes mais altas. Para além disso, todas elas apresentam uma configuração simétrica.

A Figura 63, representa a interface de utilizador “*envolventes_espectrais*”, que nos permite escutar o resultado sonoro da reprodução da Ilusão Sonora da Escala de Shepard realizada a partir da reconstituição em *Pure Data* das envolventes espectrais utilizadas por outros autores.

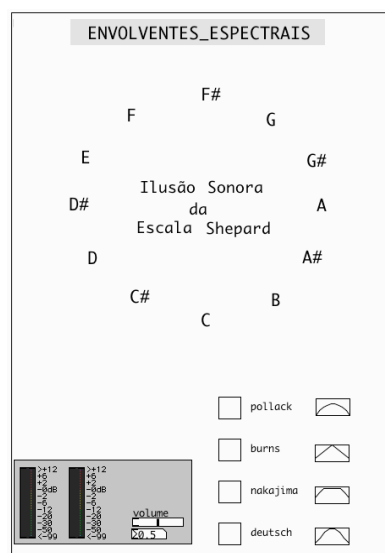


Figura 63. Interface de utilizador “*envolventes_espectrais*” usada para escutar o resultado sonoro da reprodução da Ilusão Sonora da Escala de Shepard realizada através de envolventes espectrais de outros autores.³⁶

³⁶ A interface de utilizador pode ser escutada no DVD em anexo, no directório *patches*, com o nome *16.envolventes_espectrais.pd*

Não faz parte desta dissertação, no que respeita à reprodução da Ilusão Sonora da Escala de Shepard, realizar um estudo estatístico sobre a avaliação sonora das envolventes espectrais dos outros autores aplicadas ao Chromasom. Esse estudo, será realizado no âmbito de uma experiência auditivo-musical e unicamente em relação à envolvente espectral alternativa que vamos propor. No entanto, não deixamos de mencionar, que após escutarmos o resultado sonoro das envolventes espectrais dos outros autores (através da interface de utilizador “*envolventes_espectrais*”, que se encontra representada na Figura 63) foi possível reproduzir a Ilusão Sonora da Escala de Shepard.

Na próxima secção da dissertação vamos apresentar uma envolvente espectral alternativa, tanto quanto sabemos nunca antes foi utilizada para reproduzir a Ilusão Sonora da Escala de Shepard.

IX.4 Determinação da Envolvente Espectral Alternativa

A Envolvente espectral alternativa utilizada na construção do Chromasom baseia-se nas curvas de igual percepção de intensidade sonora (*Equal-loudness Curves*).

IX.4.1 Curvas de Igual Percepção de Intensidade Sonora

As curvas de igual percepção de intensidade sonora apresentadas neste Subcapítulo estão relacionadas unicamente com sons puros.

A percepção de intensidade sonora (*Loudness*) depende simultaneamente da intensidade (i.e., do nível de pressão sonora) e do espectro sonoro do estímulo (Loy 2006, pp. 166-167; Suzuki e Takeshima 2004, p. 918).

O ser humano não ouve todas as frequências com a mesma intensidade sonora. Por exemplo, sons com conteúdo espectral entre os 1000 Hz e os 3500 Hz são ouvidos com maior intensidade sonora do que sons com igual intensidade, mas com o espectro sonoro fora desse intervalo (Loy 2006, p. 167).

É possível através das curvas de igual percepção de intensidade sonora relacionar a percepção de intensidade com a altura dos sons (i.e., com as frequências).

Os primeiros estudos de medição de níveis de igual percepção de intensidade sonora foram relatados por Kingsbury em 1927. As medições foram obtidas a partir de difusão sonora monofónica (Suzuki e Takeshima 2004, p. 918).

O primeiro estudo realizado em espaço aberto e obtido sob condições de difusão sonora estereofônica foi realizado por Fletcher e Munson em 1933 (Loy 2006, p. 168; Suzuki e Takeshima 2004, p. 918).

Desde então, têm surgido vários estudos dos quais se destaca o estudo realizado por Robinson e Dadson de 1956 (Suzuki e Takeshima 2004, p. 918). Este estudo foi adotado até 2003, como padrão internacional de medição de igual percepção de intensidade sonora para sons puros escutados em espaço aberto (ISO/R 226, 1961; ISO 226, 1987).

Suzuki e Takeshima (2004) realizaram um estudo comparativo sobre a medição de igual percepção de intensidade sonora e estabeleceram uma nova série de curvas. Consequentemente, em 2003 propuseram as suas curvas à *International Organization of Standardization* (ISO), tendo sido, desde então reconhecidas como o novo padrão internacional de medição de igual percepção de intensidade sonora para sons puros escutados em espaço aberto (BS ISO 226:2003).

Os estudos para determinar as curvas de igual percepção de intensidade sonora basearam-se no seguinte procedimento: perguntavam às pessoas para indicarem quando é que dois sons puros com frequências diferentes (sendo um deles o som de referência de 1000 Hz) eram percebidos com a mesma intensidade sonora (Scharine, Cave e Letowski 2009, p. 407). Através deste processo construíram gráficos onde se observam as curvas de igual percepção de intensidade sonora. Estas curvas não são mais do que uma média dos resultados obtidos das várias pessoas que se submeteram às experiências.

Através delas, é possível constatar que o ser humano é mais sensível a frequências que se situam entre os 1000 Hz e os 5000 Hz (Mathews 1999, p. 72). Fora desse intervalo espectral, os sons precisam de ter maior intensidade sonora para serem detectados (Plack 2004, p. 12).

A unidade de medida utilizada para relacionar a percepção de intensidade sonora e a altura dos sons chama-se *phon*. O *phon* é a unidade de medida usada nas curvas de igual percepção de intensidade sonora (Loy 2006, p. 167). Um som tem um nível de percepção de intensidade sonora de N *phons*, se produzir a mesma percepção de intensidade sonora de um som puro de 1000 Hz de N dB (Scharine, Cave e Letowski 2009, p. 407). Por exemplo, na curva de igual percepção de intensidade sonora de 40 *phons* de Fletcher e Munson, uma frequência de 100 Hz precisa de uma intensidade sonora de 60 dB para produzir a mesma percepção de intensidade sonora de uma frequência de 1000 Hz a 40 dB.

Suzuki e Takeshima (2004, pp. 928-929) compararam as suas curvas de igual percepção de intensidade sonora com as curvas de Robinson e Dadson e observaram um desvio acentuado entre os dois conjuntos de curvas de 20 a 80 *phons*. Este desvio é especialmente evidente nas regiões abaixo de 1 kHz como se pode verificar através da Figura 64.

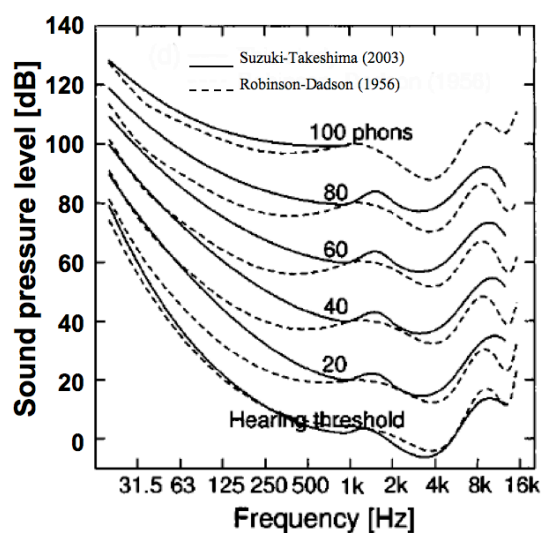


Figura 64. Comparação entre as novas curvas propostas por Suzuki e Takeshima e as curvas de Robinson e Dadson. Adap. de Suzuki e Takeshima (2004, p. 928).

Compararam também as suas curvas com as curvas publicadas por Fletcher e Munson e concluíram, que apesar de existirem algumas diferenças era importante notar que os dois conjuntos de curvas de 40 *phons* eram praticamente semelhantes ao longo do eixo das frequências (Figura 65).

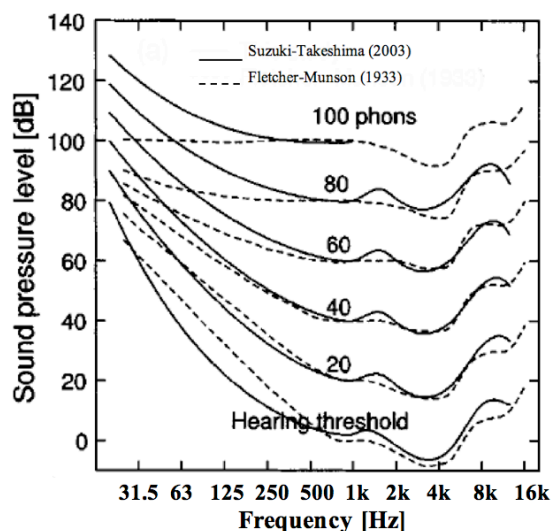
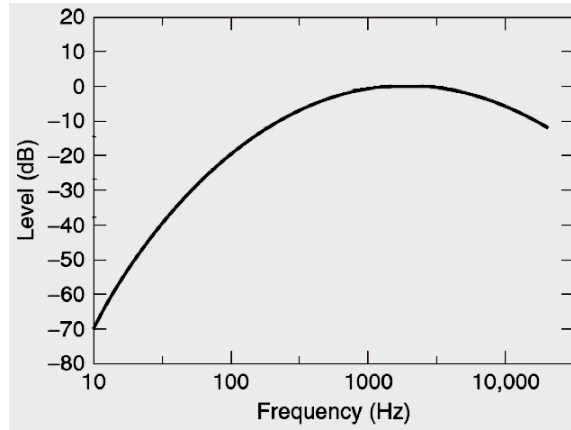


Figura 65. Comparação entre as novas curvas propostas por Suzuki e Takeshima e as curvas de Fletcher e Munson. Adap. de Suzuki e Takeshima (2004, p. 928).

Foram extraídas das curvas de Fletcher e Munson três tipos diferentes de filtros utilizados na medição da intensidade sonora do ruído. Referimo-nos concretamente aos Filtros do tipo A, B e C. Estes Filtros não são mais do que versões simplificadas das funções inversas das curvas de 40, 70 e 100 *phons* (Loy 2006, p. 169).

Segundo Wong (2007, p. 458) o Filtro C é usado sobretudo na medição das emissões acústicas produzidas por máquinas, enquanto que em relação ao Filtro B, o mesmo autor, sem adiantar os motivos, afirma que este, tem vindo a ser omitido das novas normas de medição de intensidade sonora. Por outro lado, o Filtro A, é de todos os Filtros mencionados, o que melhor descreve a sensibilidade auditiva dos seres humanos em relação às frequências. Por essa razão, é utilizado, quer no prognóstico da perda auditiva, quer na determinação dos efeitos do ruído sobre o nosso sistema auditivo (Wong, 2007, pp. 457-458). Assim, a envolvente espectral alternativa proposta nesta secção vai basear-se no Filtro A. Este Filtro, pretende simular a resposta do ouvido humano ao longo da curva de igual percepção de intensidade sonora de 40 *phons* (Wong 2007, p. 457) (Figura 66).



66. Filtro A. Nos 1000 Hz (frequência de referência) o nível de pressão sonora é igual a zero dB. Adap. de Wong (2007, p. 458).

Wong (2007, p. 459) apresenta uma Equação Matemática (Equação 10) que permite reconstruir o Filtro A, e conseqüentemente determinar a igual percepção de intensidade sonora para qualquer frequência introduzida na sua variável. Ou seja, qualquer frequência introduzida na variável da Equação 10, será escutada com a mesma intensidade sonora de um som puro de 1000 Hz (frequência de referência).

$$W_A(f) = 20 \log \left[\frac{f_4^2 f^4}{(f^2 + f_1^2)(f^2 + f_2^2)^{\frac{1}{2}}(f^2 + f_3^2)^{\frac{1}{2}}(f^2 + f_4^2)} \right] - W_{A1000} \quad (10)$$

Onde:

$W_{A1000} = -2$ dB, é a constante de normalização (i.e., a atenuação necessária para se obter a ponderação de zero decibéis em 1000 Hz);

f representa a frequência da variável em Hz;

$$f_1 = 20.6 \text{ Hz};$$

$$f_2 = 107.7 \text{ Hz};$$

$$f_3 = 737.9 \text{ Hz};$$

$$f_4 = 12194 \text{ Hz}.$$

Todos os elementos da Equação 10 (à excepção da variável) são constantes e os seus valores foram fornecidos pelo próprio autor (Wong 2007).

A implementação da Equação 10 e a extracção automática das listas das amplitudes individuais encontram-se representadas na Figura 67.

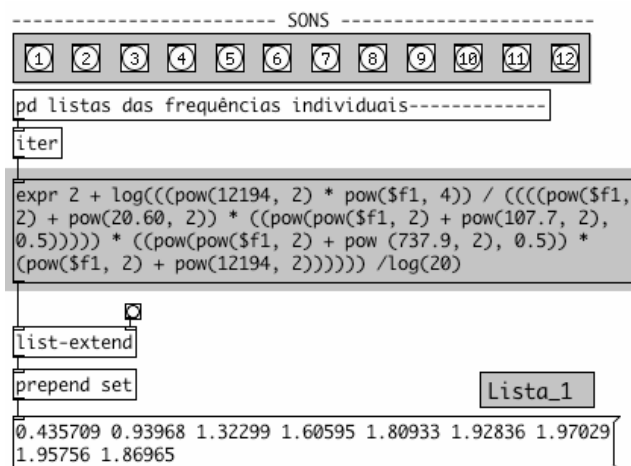


Figura 67. Implementação e extracção automática das listas das amplitudes individuais obtidas a partir da Equação 10.³⁷

A Figura 68, representa a imagem da envolvente espectral alternativa obtida para o Chromasom.

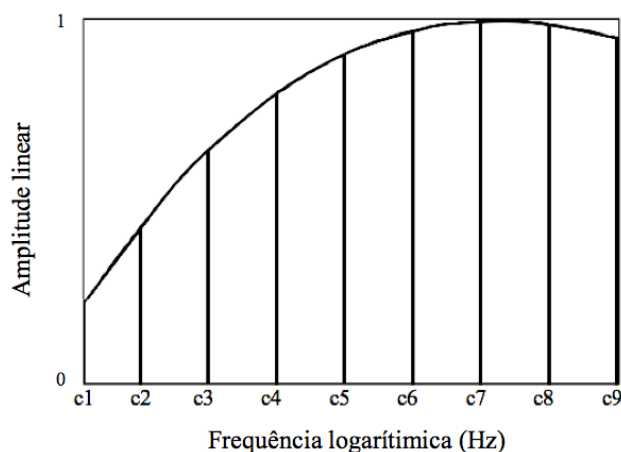


Figura 68. Representação gráfica da envolvente espectral alternativa. As linhas verticais representam nove componentes separadas por intervalos de oitava.³⁸

³⁷ A implementação da Equação 10 pode ser consultada no DVD em anexo, no directório *patches*, com o nome *17.equacao_10_envolvente_spectral_alternativa.pd*

³⁸ Para reduzir o processamento computacional convertimos os valores inicialmente obtidos, compreendidos entre zero e dois numa escala linear de zero a um.

Se invertermos a curva de 40 *phons* de Fletcher e Munson ou de Suzuki e Takeshima (2004, p. 928) e se as compararmos com o Filtro A, ou com a envolvente espectral alternativa do Chromasom verificamos que todas elas apresentam uma configuração bastante semelhante (Figura 69).

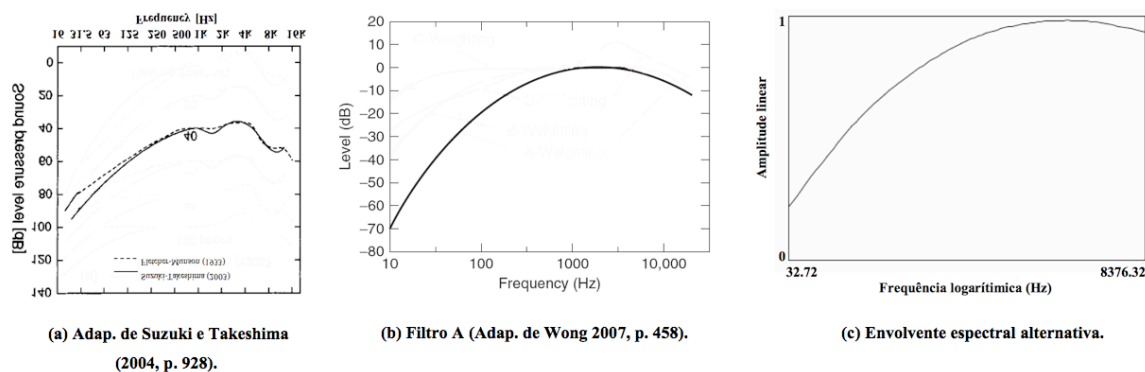


Figura 69. Comparação entre as curvas invertidas de 40 *phons* de Fletcher e Munson e de Suzuki e Takeshima (a) com a curva do Filtro A (b) e a envolvente espectral alternativa obtida para o Chromasom (c). Para que todas as curvas apresentassem a mesma posição, de maneira a facilitar a comparação, invertemos a imagem do canto esquerdo.

É importante realçar que ao contrário das envolventes espectrais anteriormente apresentadas (Subcapítulo IX.2 – Envolventes Espectrais Utilizadas por Outros Autores Aplicadas ao Chromasom), na envolvente espectral alternativa as amplitudes mais altas encontram-se sempre nas componentes mais à direita do espectro sonoro. Para além disso, apresenta uma configuração não simétrica.

Os valores das amplitudes individuais extraídos a partir do procedimento apresentado na Figura 67, estão representados na Figura 70 e na Tabela 7.

A-weighting1 0 0.435709 0.93968 1.32299 1.60595 1.80933 1.92836 1.97029 1.95756 1.86965	A-weighting7 0 0.704315 1.1453 1.47532 1.71737 1.87989 1.95715 1.97058 1.92666 1.77749
A-weighting2 0 0.483014 0.97596 1.35004 1.62582 1.82263 1.93442 1.9712 1.95383 1.85694	A-weighting8 0 0.745721 1.17681 1.4985 1.73409 1.88947 1.96032 1.9694 1.91923 1.75828
A-weighting3 0 0.529268 1.01143 1.37641 1.64518 1.83532 1.93995 1.97176 1.94959 1.84321	A-weighting9 0 0.786233 1.20754 1.5211 1.75027 1.89844 1.96308 1.96786 1.91104 1.73795
A-weighting4 0 0.574496 1.04609 1.4021 1.66401 1.84739 1.94497 1.97198 1.94479 1.82842	A-weighting10 0 0.825868 1.23752 1.54313 1.7659 1.9068 1.96544 1.96592 1.90202 1.71651
A-weighting5 0 0.618727 1.07995 1.42713 1.68232 1.85885 1.94949 1.97186 1.9394 1.81255	A-weighting11 0 0.864645 1.26674 1.56461 1.78095 1.91457 1.96742 1.96357 1.89215 1.69398
A-weighting6 0 0.661992 1.11302 1.45154 1.70011 1.86968 1.95354 1.9714 1.93337 1.79558	A-weighting12 0 0.902578 1.29522 1.58555 1.79544 1.92175 1.96903 1.96079 1.88138 1.67036

Figura 70. Listas dos valores obtidos através do procedimento representado na Figura 67.

Tabela 7. Valores das amplitudes individuais do Chromasom obtidos a partir da Equação 10. Verifica-se que as amplitudes individuais aumentam gradualmente desde a primeira até à sétima componente e diminuem desde à sétima até à nona componente de cada um dos sons.

Número de Sons (s)	Número de componentes e valores das amplitudes individuais do Cromasom								
	c1	c2	c3	c4	c5	c6	c7	c8	c9
s1	0.435709	0.93968	1.32299	1.60595	1.80933	1.92836	1.97029	1.95756	1.86965
s2	0.483014	0.97596	1.35004	1.62582	1.82263	1.93442	1.9712	1.95383	1.85694
s3	0.529268	1.01143	1.37641	1.64518	1.83532	1.93995	1.97176	1.94959	1.84321
s4	0.574496	1.04609	1.4021	1.66401	1.84739	1.94497	1.97198	1.94479	1.82842
s5	0.618727	1.07995	1.42713	1.68232	1.85885	1.94949	1.97186	1.9394	1.81255
s6	0.661992	1.11302	1.45154	1.70011	1.86968	1.95354	1.9714	1.93337	1.79558
s7	0.704315	1.1453	1.47532	1.71737	1.87989	1.95715	1.97058	1.92666	1.77749
s8	0.745721	1.17681	1.4985	1.73409	1.88947	1.96032	1.9694	1.91923	1.75828
s9	0.786233	1.20754	1.5211	1.75027	1.89844	1.96308	1.96786	1.91104	1.73795
s10	0.825868	1.23752	1.54313	1.7659	1.9068	1.96544	1.96592	1.90202	1.71651
s11	0.864645	1.26674	1.56461	1.78095	1.91457	1.96742	1.96357	1.89215	1.69398
s12	0.902578	1.29522	1.58555	1.79544	1.92175	1.96903	1.96079	1.88138	1.67036

A partir deste momento, a envolvente espectral alternativa passará a chamar-se Envolvente Espectral Alternativa Auto-Generativa (EEAAG). É auto-generativa, porque permite em tempo real, através de uma aplicação com uma interface de utilizador preservar a igual percepção de intensidade sonora de qualquer frequência introduzida na sua variável. Ou seja, qualquer frequência, cuja intensidade sonora caia ao longo da curva de 40 *phons*, será ouvida com a mesma intensidade sonora de um som puro de 1000 Hz (frequência de referência).³⁹

³⁹ A EEAAG pode ser consultada através da aplicação da interface de utilizador que se encontra no DVD em anexo, no directório *patches*, com o nome *18.EEAAG.pd*

O resultado sonoro final do Chromasom pode ser escutado através da interface de utilizador que se encontra representada na Figura 71.

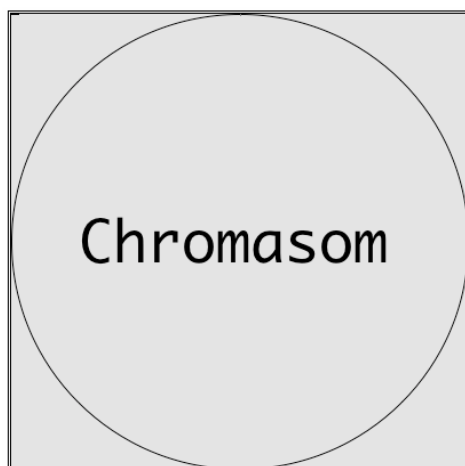


Figura 71. Interface de utilizador “*Chromasom*”.⁴⁰

No Capítulo X, desta dissertação vamos submeter o Chromasom a duas séries de exercícios musicais, e posteriormente, a uma experiência auditivo-musical com o objectivo de verificar, se o efeito da Ilusão Sonora da Escala de Shepard, ou seja, se a sensação de perpetuidade sonora se mantém, quer durante a reprodução dos exercícios, quer na experiência auditivo-musical.

IX.5 Exercícios Musicais Aplicados ao Chromasom

Os exercícios musicais aplicados ao Chromasom encontram-se distribuídos por duas séries de exercícios e por uma experiência auditivo-musical. As duas séries de exercícios representam uma etapa intermédia na construção do padrão sonoro final. Por essa razão, só serão submetidas a uma auto-avaliação. Só o padrão sonoro final (uma melodia constituída por uma sequência de saltos intervalares não regulares) é que será escrutinado, através de uma avaliação externa e de um estudo estatístico. A avaliação

⁴⁰ O Chromasom pode ser escutado através da interface de utilizador que se encontra no DVD em anexo, no directório *patches*, com o nome *19.Chromasom.pd*

externa vai realizar-se no âmbito de uma experiência auditivo-musical. Essa experiência vai ser discutida no Subcapítulo X.4, desta dissertação.

IX.5.1 Primeira Série de Exercícios Musicais

Aplicaram-se ao Chromasom duas escalas musicais divididas em 24 e 48 partes iguais (i.e., escalas de 1/4 e 1/8 de tom). Calcularam-se as frequências das notas musicais das respectivas escalas, através da Equação 11 (Pierce 1999, p. 171).

$$2^{(n/12)} * f0, \quad (11)$$

Onde:

- n é o número da nota musical da escala;
- $f0$ representa a frequência fundamental da primeira nota musical da escala;

As equivalências de oitava calculadas a partir da frequência fundamental de cada som foram determinadas através da Equação 12.

$$2^i * f0, \quad (12)$$

Onde:

- i representa o número da oitava.

A Equação 11, determina as frequências fundamentais das escalas e a Equação 12 as equivalências de oitava calculadas a partir das frequências fundamentais de cada som.

A implementação das respectivas Equações encontra-se representada na Figura 72.

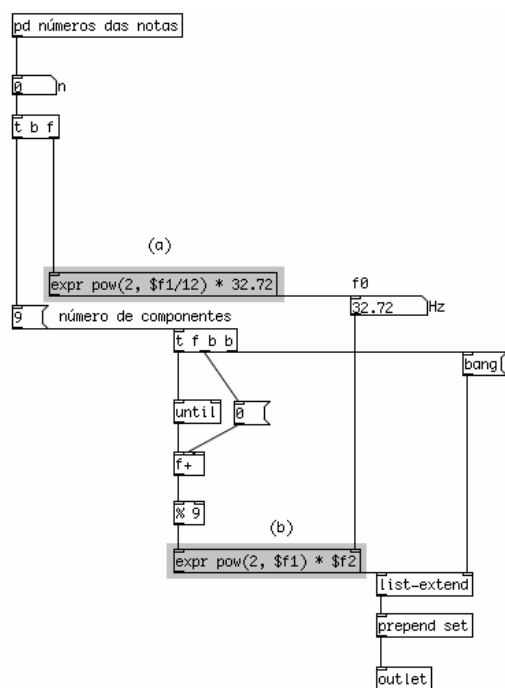


Figura 72. Implementação da Equação 11 (a) e Equação 12 (b).

Os procedimentos utilizados para a obtenção automática das frequências individuais, sob a forma de listas encontram-se representados na Figura 73.

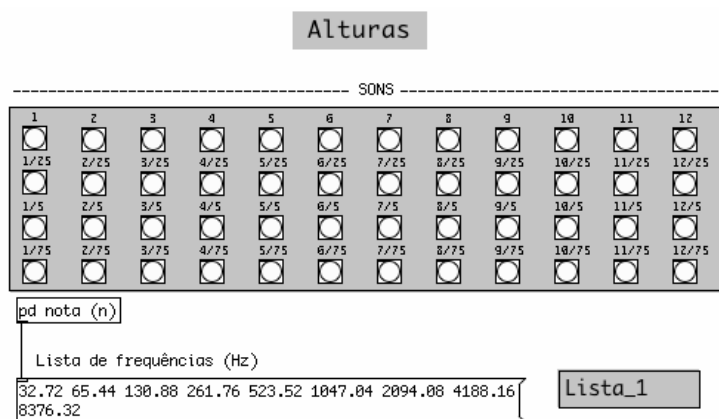


Figura 73. Determinação automática das frequências individuais das escalas musicais divididas em 24 e 48 partes iguais.⁴¹

⁴¹ A determinação automática das frequências das escalas musicais divididas em 24 e 48 partes iguais, pode ser consultada no DVD em anexo, no directório *patches*, com o nome *20.frequencias_escalas_24_e_48_partes.pd*

Após determinadas as frequências individuais, que constituem as notas musicais das respectivas escalas foram calculadas as amplitudes individuais, a partir dos procedimentos representados na Figura 74.

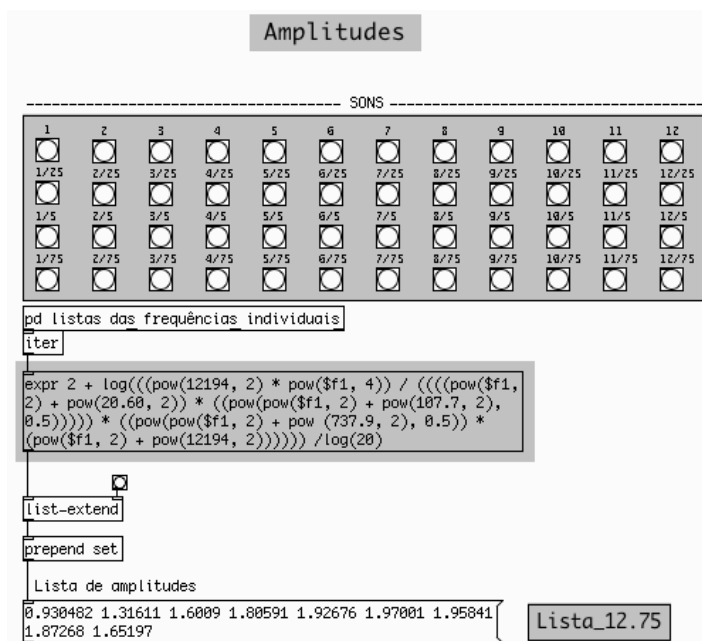


Figura 74. Determinação automática das amplitudes individuais das escalas musicais divididas em 24 e 48 partes iguais, obtidas a partir da Equação 10.⁴²

IX.5.2 Segunda Série de Exercícios Musicais

Aplicou-se ao Chromasom uma segunda série de exercícios musicais. Os exercícios são:

- progressão melódica de um intervalo de terceira maior descendente por meio-tom cromático ascendente (e.g., E-C, F-C#, F#-D, e assim sucessivamente até D#-B);
- progressão melódica de uma tríade maior descendente por meio-tom cromático descendente (e.g., G-E-C, Gb-Eb-Cb, F-D-Bb, e assim sucessivamente até Ab-F-Db);

⁴² A determinação automática das amplitudes individuais das escalas musicais divididas em 24 e 48 partes iguais obtidas através da Equação 10, pode ser consultada no DVD em anexo, no directório *patches*, com o nome *21.amplitudes_escalas_24_e_48_partes.pd*

- progressão melódica de um harpejo descendente de um acorde maior com sétima maior por meio-tom cromático descendente (e.g., B-G-E-C, B \flat -G \flat -E \flat -C \flat , A-F-D-B \flat , e assim sucessivamente até C-Ab-F-Db);
- progressão harmónica de um acorde maior com sétima maior por meio-tom cromático ascendente (e.g., C E G B, C \sharp F G \sharp C, D F \sharp A C \sharp , e assim sucessivamente até B D \sharp F \sharp A \sharp).

Os exercícios de ambas as séries podem ser escutados através da interface de utilizador “*series_exercicios*”, que se encontra representada na Figura 75.

Figura 75. Interface de utilizador “*series_exercicios*”. Esta interface permite escutar os resultados sonoros obtidos a partir das duas séries de exercícios musicais aplicados ao Chromasom.⁴³

⁴³ A interface de utilizador pode ser consultada no DVD em anexo, no directório *patches*, com o nome *22_series_exercicios.pd*

IX.5.3 Percepção Auditiva das Duas Séries de Exercícios Musicais Aplicadas ao Chromasom

No âmbito da auto-avaliação realizada verificou-se, que a sensação de perpetuidade sonora, não só se manteve em todos os exercícios, como se tornou mais evidente à medida que a complexidade do padrão sonoro aumentou. Ou seja, desde o padrão sonoro constituído por um intervalo de terceira maior (duas notas musicais) até aos harpejos de acordes maiores com sétima maior (quatro notas musicais).

Para além disso, a sensação de perpetuidade sonora também se manteve durante a repetição da progressão harmónica de um acorde maior com sétima maior por meio-tom cromático ascendente.

Para tornar ainda mais complexos os padrões sonoros previamente utilizados nas séries de exercícios apresentadas (i.e., aumentar ainda mais o número de notas musicais) e consequentemente realizar uma avaliação externa ao Chromasom (através da colaboração de ouvintes, do preenchimento de um formulário e de um estudo estatístico) realizámos uma experiência auditivo-musical, que irá ser devidamente desenvolvida na próxima secção da dissertação.

IX.5.4 Experiência Auditivo-Musical

Nesta secção, o Chromasom foi aplicado a dois exercícios musicais com o propósito de verificar, em primeiro lugar, se a Ilusão Sonora da Escala de Shepard é reproduzida com sucesso e em segundo lugar, se o efeito sonoro da ilusão, ou seja, se a sensação de perpetuidade sonora se mantém durante a repetição de uma melodia constituída por uma sequência de saltos intervalares não regulares.

A melodia mais mediática reproduzida através do Som de Shepard intitula-se *The Endless Stairs* do jogo Super Mário 64, Nintendo, composta por Koji Kondo (Figura 76).



Figura 76. *The Endless Stairs*.⁴⁴

O padrão sonoro básico desta melodia é constituído por uma sequência restrita de saltos intervalares regulares (trítonos separados por meio-tom, seguidos por mais trítonos) por exemplo: B-F, F#-C. Pelo contrário, a melodia que propomos nesta dissertação, distingue-se da anterior, pelo facto, dos saltos intervalares não serem regulares, como se pode verificar através do seu pentagrama representado na Figura 77.



Figura 77. Pentagrama da melodia. A melodia é constituída por dez notas musicais distribuídas ao longo de dois compassos (padrão sonoro). De dois em dois compassos a melodia sobe um semi-tom e é repetida durante 12 vezes.

As notas musicais estabelecem, sequencialmente entre si, intervalos não regulares, por exemplo: o primeiro, é de segunda menor (E-F); o segundo, de segunda maior (F-G); o terceiro, de terceira menor descendente (G-E); o quarto, de terceira maior descendente (E-C); o quinto, de terceira menor descendente (C-A); o sexto, de segunda maior (A-B); o sétimo, de segunda menor (B-C); o oitavo, de quarta perfeita descendente (C-G) e o nono, de terceira menor descendente (G-E).

A melodia que propomos apresenta a configuração de um ciclo estranho (Hofstadter 1979, p. 18). Se seguirmos, passo a passo as várias etapas do sistema (i.e., se acompanharmos as notas musicais que sobem meio-tom, de dois em dois compassos) deparamo-nos de repente, que regressamos sempre ao mesmo ponto de partida. Referimo-

⁴⁴ <http://www.ninsheetm.us/browse/series/SuperMario>

nos concretamente, à passagem de meio-tom cromático ascendente, entre a última nota do último compasso e a primeira nota do compasso inicial.

IX.5.4.1 Método

Na experiência foram usados 20 sujeitos constituídos exclusivamente por alunos das Classes de Instrumento e de Composição da ESMAE-IPP, Porto – Portugal.

Do grupo de sujeitos fazem parte 14 instrumentistas, seis compositores, 12 dos quais são do sexo masculino, oito do sexo feminino, com uma média de idades de 20 anos (entre 18 e os 30 anos) e uma média de 14 anos de treino musical (entre os seis e os 20 anos).

Foram utilizados como equipamento um computador portátil, uma aplicação com uma interface de utilizador construída em *Pure Data*, para tocar os exercícios musicais e um Combo Roland AC-60 para difundir o som.⁴⁵

Os exercícios foram difundidos numa sala de ensaio de orquestra a uma intensidade moderada e de uma só vez para todos os presentes.

Foi entregue um formulário a cada sujeito. Antes do seu preenchimento, os sujeitos tiveram a oportunidade de ler informação sobre a Ilusão Sonora da Escala de Shepard, ouvir uma explicação oral sobre o tema em questão e escutar o som original da respectiva ilusão (o som de referência). Com esta informação, a concentração dos ouvintes foi dirigida para o aspecto mais importante que pretendíamos testar. Referimo-nos concretamente, à sensação de perpetuidade sonora da ilusão. Por outro lado, o facto de os ouvintes possuírem informação genérica sobre a ilusão, nomeadamente o nome, o som original e a sensação auditiva que poderia provocar, não altera a sua audição. Sabemos que as ilusões persistem mesmo quando temos conhecimento do seu mecanismo de funcionamento e da expectativa que podem gerar. Esta situação pode ser corroborada através da ilusão visual da imagem ambígua (Figura 17). Apesar de sabermos, que a ilusão contém a palavra *Liar*, nunca deixamos, no entanto de percepcionar o perfil de um rosto humano com óculos, sempre que para ela olhamos. Ou, no exemplo sonoro de Bregman designado por *stream segregation in a cycle of six tones*. Apesar de sabermos

⁴⁵ A aplicação correspondente pode ser descarregada e/ou os exemplos áudio escutados em <http://sites.google.com/site/pp2007pt/uk/auditory-illusion> ou ainda, através do DVD em anexo, no directório experiência auditivo-musical, com o nome *aplicacao.pd*

que os sons (graves e agudos) são apresentados de forma alternada, quando tocados num tempo rápido, ouvimos sempre dois sons contínuos constituídos por dois grupos de registos sonoros bem distintos (um grupo de registo grave e outro de registo agudo). Ou ainda, através da *Temporal Induction of Speech: phonemic restoration* de Warren. Apesar de sabermos, antecipadamente, que foi retirado um fonema da frase e de este, ter sido substituído por um ruído, nunca deixamos de ouvir a frase intacta. Para além disso, sabemos que a percepção, em termos gerais e em termos particulares, da audição não provém somente das informações disponíveis no próprio estímulo (e.g., informação sobre a percepção da altura sonora, localização espacial, intensidade, duração, timbre, etc.) (processamento ascendente), mas também de fenómenos internos, como as expectativas e o grau de conhecimento prévio do observador (processamento descendente) (Levitin 2008 p. 34). Com este procedimento, pensamos ter aplicado o formulário de uma forma científica, rigorosa e eficaz.

Subsequentemente, foi pedido aos sujeitos que respondessem SIM ou NÃO às seguintes questões:⁴⁶

1. O exercício número um, reproduz com sucesso a Ilusão Sonora da Escala de Shepard?
2. Se no exercício número dois, o efeito da Ilusão Sonora da Escala de Shepard se mantém.

IX.5.4.2 Resultados

Depois de recolhidos os formulários e realizado o tratamento dos dados obtiveram-se os seguintes resultados em relação ao exercício número um (Tabela 8).

Tabela 8. Apresentação dos resultados obtidos no exercício número um.

Exercício Número Um		
Respostas	Sujeitos	Percentagens
SIM	17	85%
NÃO	3	15%

⁴⁶ O formulário encontra-se no anexo 4.

Constata-se, que 17 sujeitos responderam SIM, enquanto três responderam NÃO. Significa, que 85 por cento dos sujeitos inquiridos considerou, que o Chromasom reproduziu com sucesso a Ilusão Sonora da Escala de Shepard, enquanto 15 por cento respondeu negativamente.

Os resultados obtidos relativos ao exercício número dois estão apresentados na Tabela 9.

Tabela 9. Apresentação dos resultados obtidos no exercício número dois.

Exercício Número Dois		
Respostas	Sujeitos	Percentagens
SIM	15	75%
NÃO	5	25%

Constata-se, que 15 sujeitos responderam SIM, enquanto cinco responderam NÃO. Significa, que 75 por cento dos sujeitos inquiridos considerou, que a sensação de perpetuidade sonora foi mantida durante a repetição da melodia, enquanto que 25 por cento respondeu negativamente.

Em relação aos resultados globais (i.e., ao somatório dos resultados alcançados em ambos os exercícios) obteve-se uma média de 80 por cento de respostas SIM contra uma percentagem de 20 por cento de respostas NÃO.

Obteve-se um rácio de respostas SIM e NÃO na ordem de 4:1. Esta diferença significativa de rácio pode ser observada na Figura 78.

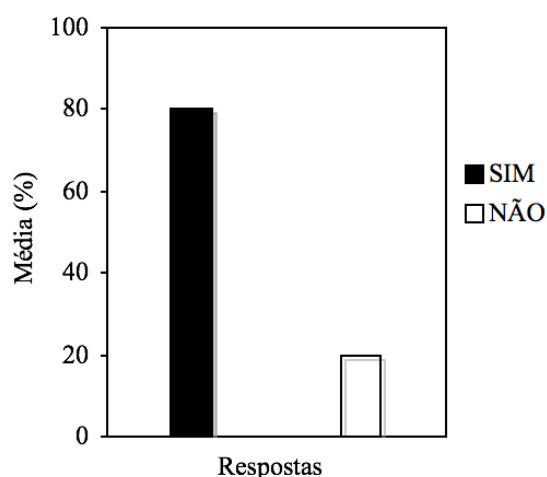


Figura 78. Média de respostas SIM e NÃO. O rácio entre as respostas SIM e Não é de 4:1.

O resultado global obtido pode ser estatisticamente interpretado da seguinte maneira: numa população de 20 sujeitos inquiridos, 4/5 (i.e., 16 sujeitos) consideraram que o Chromasom reproduziu com sucesso a Ilusão Sonora da Escala de Shepard no primeiro exercício e que a sensação de perpetuidade sonora foi mantida no segundo, enquanto 1/5 (i.e., quatro sujeitos) responderam que não.

IX.5.4.3 Discussão dos resultados

Os resultados obtidos nesta experiência demonstraram que o Chromasom foi capaz de reproduzir com sucesso a Ilusão Sonora da Escala de Shepard, mesmo quando aplicado à melodia (Figura 77).

Para além disso, através da envolvente espectral alternativa aplicada ao Chromasom, foi possível demonstrar que a envolvente espectral em forma de sino ou Curva Gaussiana usada por Shepard, não é um factor essencial para produzir com sucesso a respectiva ilusão sonora. Esta constatação já tinha sido anteriormente demonstrada por Pollack (1978, p. 203).

Foi observado uma diminuição de dez por cento nas respostas SIM entre o primeiro e o segundo exercício. Esta diminuição pode ser explicada da seguinte maneira: alguns sujeitos (25 por cento, Tabela 9) expressaram que a alteração do centro tonal ou a alteração de tonalidade que se verifica ao longo da melodia poderá enfraquecer a sensação de perpetuidade sonora. Para estes sujeitos, a sensação de perpetuidade sonora foi mais

forte no primeiro exercício, pelo facto do ciclo ser mais curto, simples e consequentemente mais fácil de acompanhar ou de prever o regresso à tónica da escala.

Este argumento é contraditório, porque quando se detecta o regresso à tónica da escala (i.e., quando se pressente um salto descendente na percepção da altura sonora entre a última nota e a tónica da escala) significa que o ciclo é quebrado e consequentemente, que a sensação de perpetuidade sonora se perdeu.

De acordo, com a maioria dos sujeitos (75 por cento, Tabela 9) a sensação de perpetuidade foi mais forte no segundo exercício, devido ao facto de este ser mais complexo, ou seja, como tem mais notas musicais, o ciclo torna-se maior e mais difícil de acompanhar auditivamente. Para além disso, como o regresso à primeira nota do primeiro compasso é imperceptível, cria a sensação da melodia subir perpetuamente.

Nas próximas secções da dissertação vamos aplicar um circuito básico de síntese por Modulação de Frequência ao Chromasom com o objectivo de testar o seu comportamento sonoro, e simultaneamente, procurar novos timbres para compor música. Para além disso, vamos também utilizá-lo, como gerador de som na concepção de um instrumento digital multimédia e em composição musical.

X APLICAÇÕES DO CHROMASOM E COMPOSIÇÃO

Para testar o comportamento sonoro e a plasticidade musical do Chromasom vamos primeiramente, aplicá-lo a um circuito básico de síntese por modulação de frequência e posteriormente, utilizá-lo como gerador de som, quer na concepção de um instrumento digital multimédia, quer em composição musical.

X.1 Aplicação do Chromasom em Síntese por Modulação de Frequência

Começamos por realizar uma breve revisão bibliográfica sobre o tema em questão, e de seguida, aplicamos a respectiva técnica ao Chromasom.

X.1.1 Modulação de Frequência

Inicialmente, a técnica de síntese por Modulação de Frequência (*Frequency Modulation* ou FM) foi aplicada às transmissões de rádio. Chowning (1973, p. 526) criou um algoritmo de FM para ser utilizado em síntese de som. O seu algoritmo foi patenteado e comercializado pela *Yamaha Corporation*, através dos sintetizadores FM conhecidos por DX7.

A vantagem da utilização da FM em relação a outras técnicas de síntese de som digital, passa pelo facto de permitir produzir, de uma maneira simples, espectros sonoros complexos e por se obter um controlo relativamente fácil sobre as frequências das componentes resultantes (Chowning 1973, p. 526).

O princípio básico de síntese por FM consiste em modular a frequência de um oscilador com outro oscilador (Farnell 2008, p. 276).

Num circuito básico de síntese por FM existem três parâmetros principais (Farnell, 2008, p. 276). São eles:

- Frequência portadora (f_p). Representa a frequência que vai ser modulada ao longo do tempo.
- Frequência moduladora (f_m). Determina a velocidade de modulação que actua sobre a frequência portadora;
- Índice de modulação (i). É a amplitude do oscilador modulador e determina a quantidade de modulação exercida sobre a frequência portadora.

Quando o índice de modulação é igual a zero ($i = 0$) não existe modulação. Obtém-se unicamente a frequência portadora (Chowning 1973, p. 527; Farnell 2008, p. 278). Quando o índice de modulação é maior do que zero ($i > 0$) surgem novas frequências (frequências laterais ou *sidebands*) com valores que se encontram, quer acima, quer abaixo do valor da frequência portadora (Chowning 1973, p. 527).

À medida que se aumenta o valor do índice de modulação, a energia retirada à frequência portadora é distribuída por um número crescente de frequências laterais, que vão surgindo (Chowning 1973, p. 527).

A complexidade do espectro está relacionado com índice de modulação. Sempre que o índice de modulação aumenta a largura de banda do espectro também aumenta (Chowning 1973, p. 529).

As frequências laterais são múltiplos inteiros da f_m (Farnell 2008, p. 278). Por exemplo, com uma $f_p = 1000$ Hz e com uma $f_m = 200$ Hz, obtêm-se as seguintes *sidebands*:

$$f_p + 1f_m \Rightarrow 1000 + 200 = 1200 \text{ Hz e } f_p - 1f_m \Rightarrow 1000 - 200 = 800 \text{ Hz;}$$

$$f_p + 2f_m \Rightarrow 1000 + 400 = 1400 \text{ Hz e } f_p - 2f_m \Rightarrow 1000 - 400 = 600 \text{ Hz;}$$

$$f_p + 3f_m \Rightarrow 1000 + 600 = 1600 \text{ Hz e } f_p - 3f_m \Rightarrow 1000 - 600 = 400 \text{ Hz,}$$

e assim sucessivamente.

A Figura 79, representa a implementação em *Pure Data* de um circuito básico de síntese por FM. Através da interface de utilizador “*FM*” podem-se observar (1) os principais parâmetros utilizados na FM, (2) o espectro da FM através do “*espectrograma_FM*” e por último (3) escutar os respectivos efeitos sonoros.

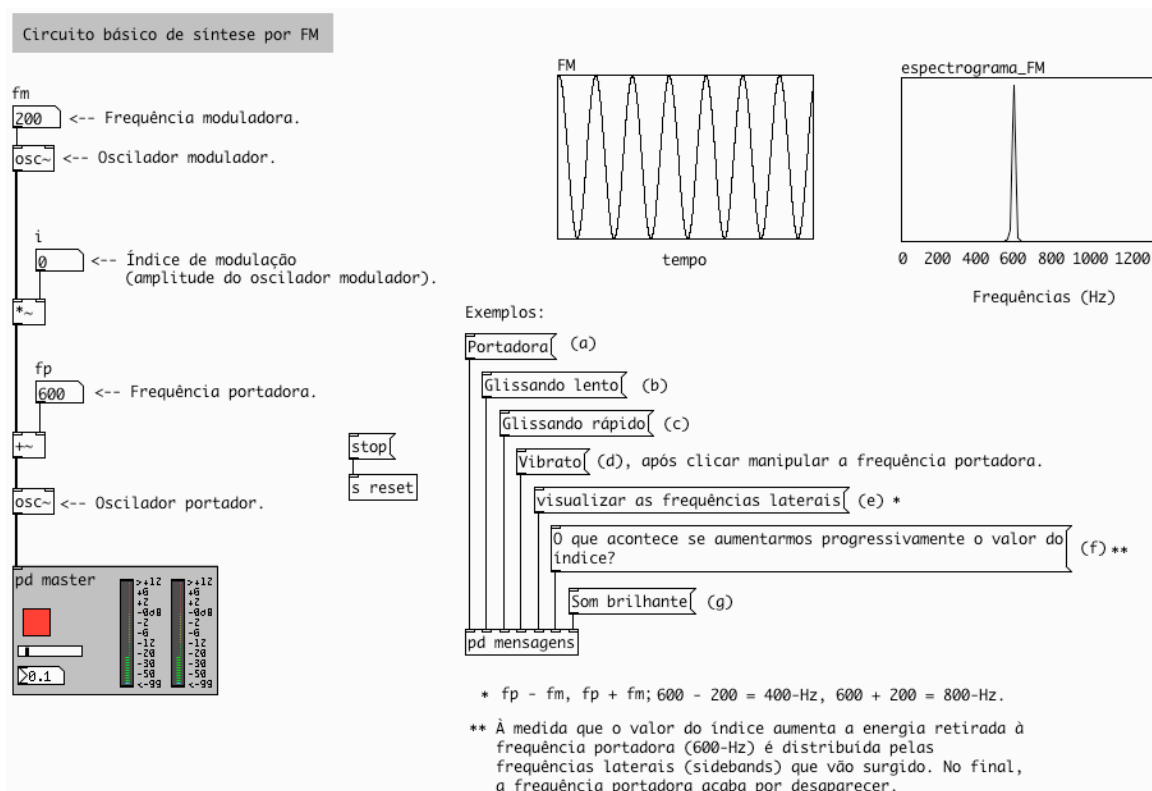


Figura 79. Interface de utilizador “FM”.⁴⁷

Através desta técnica de síntese de som, podem-se obter vários efeitos sonoros. Os mais conhecidos resultam dos seguintes procedimentos:

- com um índice de modulação e com uma frequência moduladora com valores muito baixos, por exemplo $i = 12$ e $fm = 5$ Hz, ao manipularmos a frequência portadora ficamos com a sensação de estar a ouvir uma voz humana a cantar com vibrato (Figura 79, exemplo d);
- com valores muito baixos para a frequência moduladora e com um índice de modulação elevado, por exemplo: $fm = 0.03$ Hz e $i = 200$, a frequência portadora (600 Hz) oscila lentamente entre os 400 Hz e os 800 Hz ($fp - i$, $fp + i$) originando um glissando ascendente e descendente da fp (Figura 79, exemplos b e c);
- se aumentarmos progressivamente o valor do índice de modulação fazemos com que o som se torne timbricamente, cada vez mais brilhante (Farnell 2008, p. 277) (Figura 79, exemplo g).

⁴⁷ A interface de utilizador pode ser consultada através do DVD em anexo, no directório *patches*, com o nome *23.FM.pd*

X.1.2 Um Chromasom a Modular outro Chromasom

O circuito básico de síntese por FM aplicado ao Chromasom consiste em modular um Chromasom com outro Chromasom. Estamos perante uma modulação complexa, no sentido, em que temos várias frequências portadoras (i.e., as componentes do Chromasom portador) a serem moduladas por um Chromasom (i.e., pelo Chromasom modulador) (Figura 80).

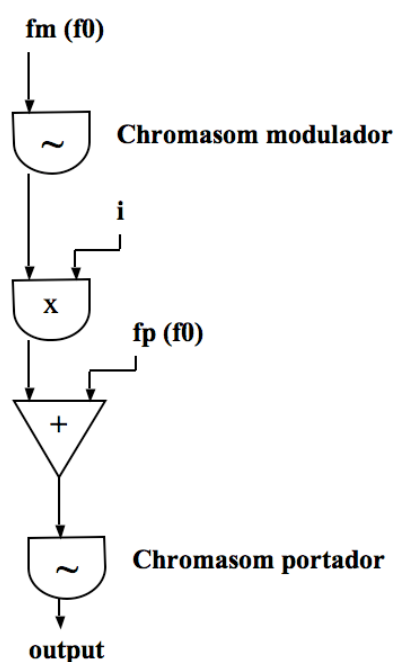


Figura 80. Representação geral de um circuito básico de síntese por FM aplicado ao Chromasom. A $f_m (f_0)$ é a frequência fundamental do Chromasom modulador e a $f_p (f_0)$ é a frequência fundamental do Chromasom portador. O índice de modulação (i) é a amplitude do Chromasom modulador e determina a quantidade de modulação exercida sobre cada uma das frequências das componentes do Chromasom portador.

Para uma visualização mais pormenorizada da estrutura interna do circuito básico de síntese por FM aplicada ao Chromasom (implementada em *Pure Data*), apresentamos a Figura 81.

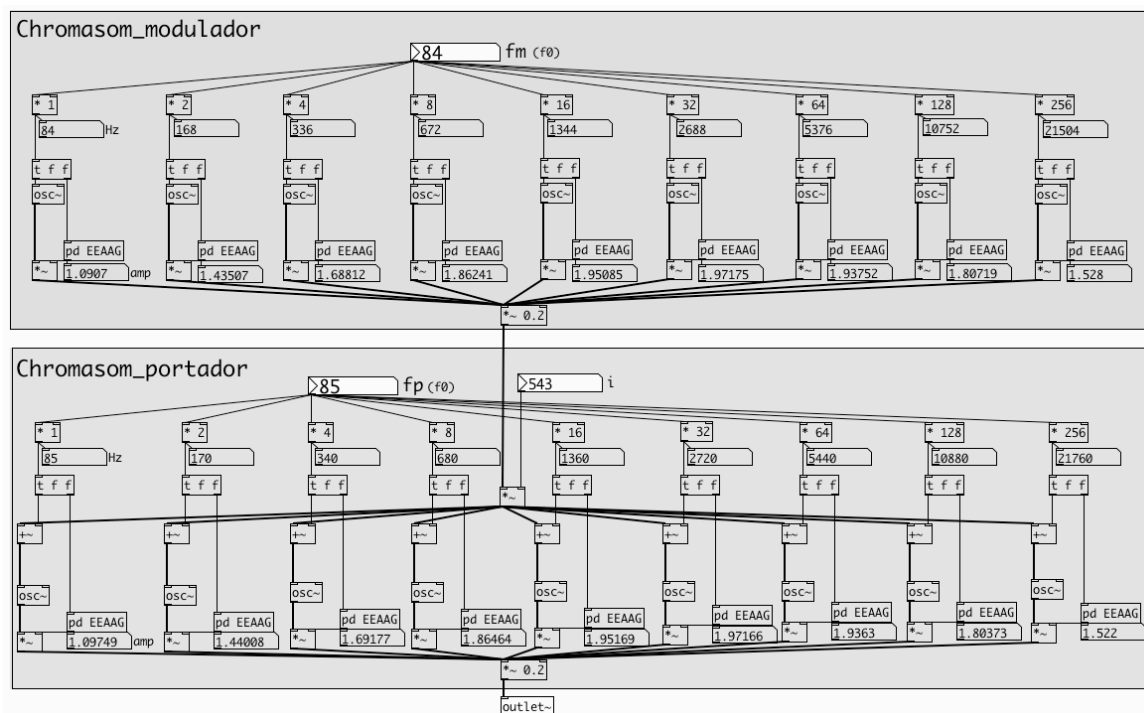


Figura 81. Implementação em *Pure Data* de um circuito básico de síntese por FM aplicado ao Chromasom. O *patch* tem a particularidade de permitir introduzir e manipular, em tempo real os valores dos principais parâmetros da FM (i.e., fm, fp e o i), de forma a explorar e produzir diferentes timbres sonoros, que resultem musicalmente interessantes, para usar em composição musical digital. É possível observar os valores das frequências (Hz) e das amplitudes lineares de cada uma das componentes, quer do Chromasom modulador, quer do Chromasom portador. As amplitudes lineares são extraídas a partir da EEAAG e variam sempre, que se alteram os valores das frequências fundamentais (f_0), quer do Chromasom modulador, quer do Chromasom portador.⁴⁸

Para escutar alguns dos resultados sonoros obtidos, utilizar a interface de utilizador designada por “*chromasom_fm_presets*” (Figura 82).

⁴⁸ O circuito básico de síntese por FM aplicado ao Chromasom pode ser consultado através da interface de utilizador que se encontra no DVD em anexo, no directório *patches*, com o nome *24.chromasom_fm.pd*

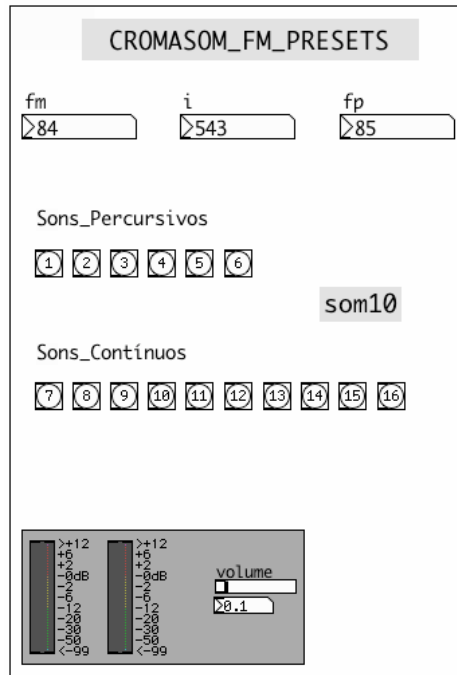


Figura 82. Interface de utilizador “*chromasom_fm_presets*”. Esta interface permite escutar os resultados sonoros obtidos a partir do circuito básico de síntese por FM aplicado ao Chromasom. É constituída por *presets* de sons percussivos e sons contínuos. Os *presets* encontram-se devidamente classificados e numerados na respectiva interface.⁴⁹

Os diferentes timbres alcançados, que nos pareceram ser musicalmente interessantes vão ser utilizados em projectos musicais. Assim, na próxima secção desta dissertação, vamos apresentar algumas aplicações musicais do Chromasom, no que respeita à concepção de um instrumento digital multimédia e à composição musical.

X.2 MuDI - Instrumento Digital Multimédia

O trabalho que vamos apresentar pertence ao contexto da interpretação musical com dispositivos electrónicos móveis e aos novos paradigmas da composição musical em tempo real.

⁴⁹ A interface de utilizador pode ser consultada no DVD em anexo, no directório *patches*, com o nome *25.chromasom_fm_presets.pd*

Em resposta a um convite para tocar música ao vivo para um filme decidimos conceber o MuDI, um instrumento digital multimédia para compor e tocar música para filmes em tempo real.

O MuDI é um instrumento multimédia, porque permite simultaneamente (1) difundir som, imagem e vídeo, (2) operar som e vídeo em tempo real, (3) gravar o som da interpretação, (4) acompanhar não só as imagens do filme, mas também a relação entre os gestos realizados pelo intérprete e o som gerado e (5) no final da interpretação obter um ficheiro audiovisual (i.e., a composição), para além de uma partitura vídeo.⁵⁰

Um dos objectivos deste trabalho, passa por tentar alcançar grande expressão musical. Desde o início, sentimos a necessidade de ter um sistema musical estável, bem como um controlo efectivo sobre o som. Como resposta a estas duas necessidades criámos (1) um processo de calibração do instrumento, que permitiu ultrapassar a falta de um ponto de referência fiável do acelerómetro (Essl e Rohs 2009, pp. 203-204) (o principal sensor utilizado para operar o MuDI) e (2) uma aplicação com uma interface de utilizador, que permite acompanhar a relação entre os gestos realizados e o som gerado.

Para além disso, o facto de termos de tocar ao vivo, em tempo real e com bastante precisão a obra musical previamente composta, sentimos a falta de ter um guia musical ou uma partitura. Assim, apresentaremos também um método para obter automaticamente uma partitura vídeo, que será útil para compor e tocar música em tempo real.

Nas próximas secções vamos abordar o trabalho relacionado com o MuDI, fazer uma descrição mais pormenorizada sobre o seu sistema, manuseamento, processo de calibração e obtenção da partitura vídeo.

X.2.1 Trabalho Relacionado

A concepção do MuDI baseia-se em trabalhos anteriores onde se exploram as possibilidades de utilização de dispositivos electrónicos móveis, por exemplo PDAs (*Personal Digital Assistants*) e telemóveis (*smartphones*) como controladores ou instrumentos musicais.

Geiger (2003, 2006), explorou a possibilidade de usar como controlador musical a tela de toque único de um PDA. Implementou uma interface de utilizador constituída por

⁵⁰ Um vídeo promocional sobre o MuDI em <https://www.youtube.com/watch?v=leh-FMwIh5g> ou no DVD em anexo, no directório MuDI, com o nome *MuDI_promoção.mov*

um sintetizador FM, uma guitarra e um conjunto de instrumentos de percussão programados em *Pure Data* (Geiger 2003, pp. 2-3; 2006, pp. 61-63). Nesse projecto, o utilizador opera a interface, quer através de uma ponteira de toque (*stylus*) ou directamente com os dedos. Para alcançar expressão musical foram utilizados gestos como pressionar, dedilhar e bater. No entanto, como a tela de toque só permitia detectar um ponto de contacto de cada vez, dificultou a obtenção de expressão musical.

Novas abordagens, como o acoplamento de unidades de sensores externas ao dispositivo manual, permitiram ao utilizador não só expandir como desenvolver um novo reportório gestual de interacção.

Seguindo esta abordagem, Tanaka (2004) expandiu o PDA através de uma unidade externa de sensores constituída por sensores de resisistência (*force sensing resistors*) para captar a força de pressão, um acelerómetro para captar a posição, a velocidade e a aceleração de movimento dos gestos e uma placa de rede Wi-Fi para transmitir áudio entre os utilizadores. Neste projecto, a expressão musical é alcançada da seguinte maneira: a intensidade da força com que o utilizador agarra o PDA transforma o brilho do som (i.e., o Timbre), o bater do ritmo re-sequencia a canção no domínio do tempo e quando os utilizadores se aproximam geograficamente uns dos outros, as partes individuais da canção, previamente seleccionadas por eles, tornam-se mais proeminentes na mistura final partilhada pelos utilizadores através da rede Wi-Fi.

Outros meios tecnológicos alternativos, como interacções baseadas em câmaras integradas de telemóveis foram testados em projectos como CaMus (2006). Em CaMus o utilizador move a câmara ao longo de uma superfície constituída por uma grelha de marcas visuais. A informação extraída a partir da posição, da distância e dos ângulos de rotação em relação às marcas visuais é posteriormente convertida em sons através de um sequenciador áudio. Contudo, como o movimento realizado ao longo da superfície promove constantes alterações no som, torna difícil manter a percepção da altura sonora desejada. Por outro lado, a fraca capacidade da câmara em captar movimentos rápidos causa atrasos na percepção auditiva, entre os gestos realizados e o som gerado. Para além disso, quer a extensão limitada da superfície da grelha, quer a curta distância necessária para realizar o reconhecimento, entre as marcas visuais e a lente da câmara, limitam o uso de gestos livres de interacção (Essl e Rohs 2009, pp. 199-202). Consequentemente, tornam difícil alcançar a expressão musical desejada.

O desenvolvimento do acelerómetro e a sua integração em dispositivos electrónicos móveis forneceram ao utilizador a possibilidade de executar gestos de

interacção em espaço aberto, ou seja, os gestos do utilizador para interagir com o dispositivo já não se encontram limitados ao tamanho de uma única superfície, como acontece na tela de toque único usada em Geiger ou na superfície em forma de grelha usada em CaMus. Para além disso, a capacidade dos acelerómetros em captar movimentos rápidos permite atenuar a percepção auditiva dos atrasos entre os gestos realizados e o som gerado. Os acelerómetros podem ser usados para realizar gestos precisos e captar movimentos rápidos (Essl e Rohs 2009, p. 204).

A partir da integração deste tipo de sensores em dispositivos electrónicos móveis comuns surgiu um novo vocabulário gestual de interacção para alcançar expressão musical de uma maneira mais simples e precisa. Por exemplo, gestos usados para tocar instrumentos musicais acústicos (percussão, chocalhos e violino) como bater, agitar ou deslizar foram usados para tocar música através de um telemóvel (Essl e Rohs, 2007).

No caso do projecto MoGMI (2008) o utilizador opera o telemóvel ao longo dos três eixos do acelerómetro. Os movimentos do utilizador são gravados, e posteriormente, convertidos em mensagens MIDI através de um gravador/reprodutor MIDI. Cada eixo do acelerómetro controla um parâmetro musical diferente, por exemplo o eixo x controla o ataque do som, o eixo y a percepção da altura sonora e o eixo z a percepção da intensidade sonora. No entanto, o uso de um gravador/reprodutor MIDI para converter os movimentos do utilizador em mensagens MIDI não permite tocar música em tempo real (Dekel e Dekel 2008, p. 7).

A integração e a combinação de ambas as tecnologias, ou seja, a tela multi-toque (que permite a detecção multi-ponto) e o acelerómetro no mesmo dispositivo electrónico móvel (e.g., no iPod touch da Apple) permitiram expandir, ainda mais as possibilidades de interacção e consequentemente alcançar expressão musical de uma maneira mais eficiente.

Estas novas possibilidades foram exploradas em projectos como MoPho (2008) uma orquestra de telemóveis e PDAs para compor e tocar música. A detecção da posição, velocidade, aceleração de movimento do dispositivo electrónico móvel através do acelerómetro e a detecção multi-ponto através da tela multi-toque são os principais meios de interacção usados em MoPho para alcançar expressão musical.

Quando interagimos com dispositivos electrónicos móveis baseados exclusivamente em telas de toque, geralmente ambas as mãos e olhos são necessários para os operar. Ou seja, uma mão para agarrar o dispositivo e a outra para o manipular. Para

além disso, olhar para a tela é indispensável, não só para ver a interface de utilizador, mas também para controlar o seu manuseamento.

Pelo contrário, operar um acelerómetro requer somente uma mão (i.e., a mão que agarra o dispositivo electrónico). Desta maneira, o utilizador não tem necessidade de olhar constantemente para o dispositivo para o manusear. Em relação ao instrumento digital multimédia proposto nesta secção da dissertação, este último aspecto é preponderante, no sentido, em que libertando o olhar do dispositivo o utilizador fica livre para acompanhar através da aplicação da interface de utilizador (Figura 84), não só as imagens do filme, mas também a relação entre os gestos realizados e o som gerado.

X.2.2 Sistema do MuDI

O sistema do MuDI baseia-se nos seguintes elementos tecnológicos:

- um dispositivo electrónico móvel (iPod touch, iOS versão 2.2.1);
- uma aplicação de dispositivo electrónico móvel (OSCRemote, versão 1.2);
- um computador portátil (Macbook, OSX 10.4.11);
- uma aplicação de computador (*Pure Data*, versão 0.39.3-extended).

Os elementos tecnológicos mencionados e a configuração do sistema estão devidamente descritos nas próximas secções.

X.2.2.1 Dispositivo Electrónico Móvel e Aplicação

Pelo facto, dos dispositivos electrónicos móveis actuais serem mais fáceis de manusear e de possuírem mais sensores integrados do que os computadores portáteis, escolhemos o iPod touch para funcionar exclusivamente como controlador musical.

O iPod touch foi lançado no mercado em Setembro de 2007, pela Apple Inc. Tem embutido uma placa de rede Wi-Fi, uma tela multi-toque e um acelerómetro para leitura da posição, velocidade e aceleração de movimentos em três dimensões espaciais.

Existem várias aplicações para o iPod touch, no entanto, a mais flexível é a OSCRemote, no sentido, em que permite editar a interface de utilizador na própria

aplicação, não sendo necessário recorrer a nenhum computador ou programa específico para tal.⁵¹

O OSCRemote é uma aplicação intuitiva que usa o protocolo de comunicação OpenSound Control (OSC) (Wright, 2005), permitindo que o iPod touch comunique e interaja dentro de um sistema de rede Wi-Fi com qualquer programa que suporte o respectivo protocolo.

Os controladores mais comuns da aplicação são:

- *Buttons*. Permitem enviar mensagens específicas quando pressionados.
- *Switches*. Ao acciona-los o utilizador pode alterar o seu estado (ON/OFF).
- *Sliders*. Quando manipulados emitem valores constantes.
- *Acelerómetros*. Detectam a posição espacial, a velocidade e a aceleração de movimento do dispositivo manual, nos eixos *x*, *y* e *z*.

A Figura 83, apresenta uma interface de utilizador criada e editada na própria aplicação. Nesta interface podem ser vistos os controladores acima mencionados. Depois de criados e editados, os controladores encontram-se de imediato disponíveis para uso do utilizador.

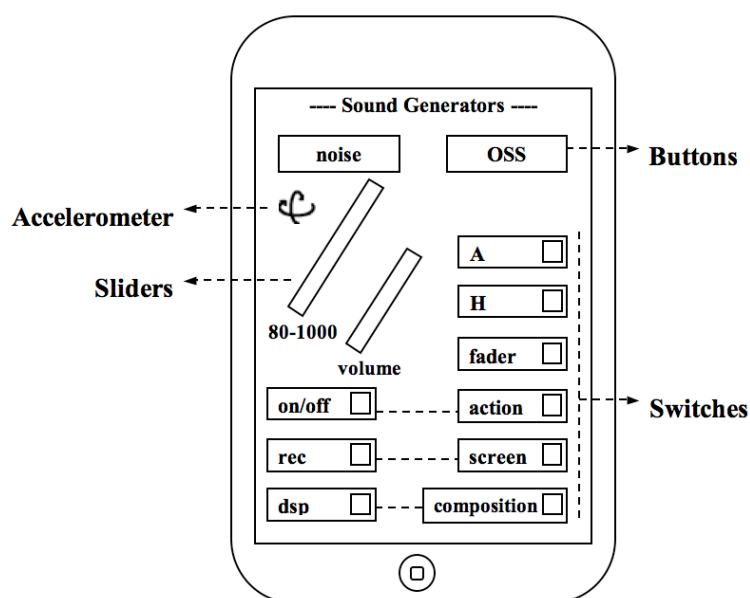


Figura 83. Exemplo de uma interface gráfica de utilizador directamente criada e editada na aplicação.

⁵¹ OSCRemote, <http://www.nr37.nl/OSCRemote/>

Este conjunto de controladores foram seleccionados para compor e tocar música digital para filmes em tempo real, pelo facto de serem intuitivos e de fácil manuseamento. Para além disso, o facto de apresentarem diferentes formas e por desempenharem diferentes funções permitem ao utilizador criar um vasto reportório gestual de interacção. Quando manuseados de forma consciente podem ser um meio eficaz de obter um controlo efectivo sobre o som e consequentemente alcançar grande expressão musical.

Oferecem ainda respostas rápidas e precisas aos gestos de interacção realizados pelo utilizador, permitindo desta maneira, interagir com precisão com os diferentes ritmos das imagens do filme projectado.

Entre os controladores apresentados, gostaríamos de realçar a versatilidade de manuseamento do acelerómetro, pelo facto de permitir realizar gestos de interacção em espaço aberto.

X.2.2.2 Computador Portátil e Aplicação

O MuDI requer pelo menos um computador portátil, porque necessita de uma placa de rede Wi-Fi, bem como uma saída de áudio e vídeo integradas.

O computador portátil foi seleccionado para realizar a produção sonora, pelo facto, de possuir uma maior capacidade de processamento de informação, armazenamento e memória interna do que qualquer PDA ou telemóvel. Para além disso, tem ao seu dispor uma grande variedade de poderosos e versáteis programas informáticos dedicados à geração de síntese de som (e.g., *Max/Msp*⁵², *Pure Data*, *SuperCollider*⁵³, *Csound*⁵⁴, *Chuck*⁵⁵).

O MuDI usa a aplicação *Pure Data extended* (*Pd-extended*), pelo facto, de esta versão incluir algumas livrarias adicionais (e.g., GEM, OSC, etc.), de maneira a permitir a reprodução vídeo e a comunicação OSC com o dispositivo electrónico móvel.

⁵² Max/Msp, <http://cycling74.com/>

⁵³ SuperCollider, <http://supercollider.sourceforge.net/>

⁵⁴ Csound, <http://www.csounds.com/>

⁵⁵ Chuck, <http://chuck.cs.princeton.edu/>

X.2.2.3 Configuração do Sistema

A configuração do sistema encoraja o uso do MuDI, porque os procedimentos de configuração são rápidos, fáceis de realizar e permanecem definitivamente gravados, não sendo necessário repetir os procedimentos de configuração iniciais.⁵⁶

Para enviar informação da aplicação do dispositivo electrónico móvel para o computador portátil (através do protocolo de comunicação OSC) é necessário que ambos partilhem o mesmo endereço e portas de comunicação do Protocolo de Internet (IP).

Quer a aplicação do dispositivo electrónico móvel, quer o *Pd-extended* permitem atribuir os mesmos endereços e designar as mesmas portas IP. No caso específico do MuDI foi escolhida a porta IP 1111.

Para estabelecer a comunicação OSC entre os controladores individuais da aplicação do dispositivo electrónico móvel e a interface de utilizador da aplicação de computador (que iremos apresentar mais à frente) temos de escolher um nome predeterminado para todas as mensagens partilhadas. Neste projecto foi escolhido o nome *“/remote”*. Assim, todos os controladores da aplicação do dispositivo electrónico móvel devem ter um único nome como *“/remote/fader1”*, *“/remote/button3”*, *“/remote/accelerometer”*, etc.

X.2.2.4 Interface de Utilizador da Aplicação de Computador

Nesta secção vamos abordar a estrutura e o modo de operação da interface de utilizador da aplicação de computador. A interface de utilizador é um *patch* de *Pd-extended* constituído por quatro módulos. Cada módulo desempenha uma função específica. Por exemplo, o Módulo Comunicação realiza a recepção, encaminhamento e tratamento individual de toda a informação transmitida pela aplicação de controlo. O Módulo Som está exclusivamente dedicado à geração de som e à gravação áudio da composição musical directamente no disco rígido do computador. O Módulo GUI, através dos objectos *Pd-extended* designados por *Graphical Users Interfaces* ou GUI, permite visualizar a calibração do sistema, os pontos de referência e a relação entre os gestos

⁵⁶ A única excepção acontece quando mudamos de rede Wi-Fi. Nesse caso, é necessário introduzir o novo endereço IP na aplicação do dispositivo electrónico móvel.

realizados e o som gerado. Por último, o Módulo Vídeo está dedicado ao visionamento das imagens do filme e à reprodução musical em sincronização com o filme projectado (Figura 84).

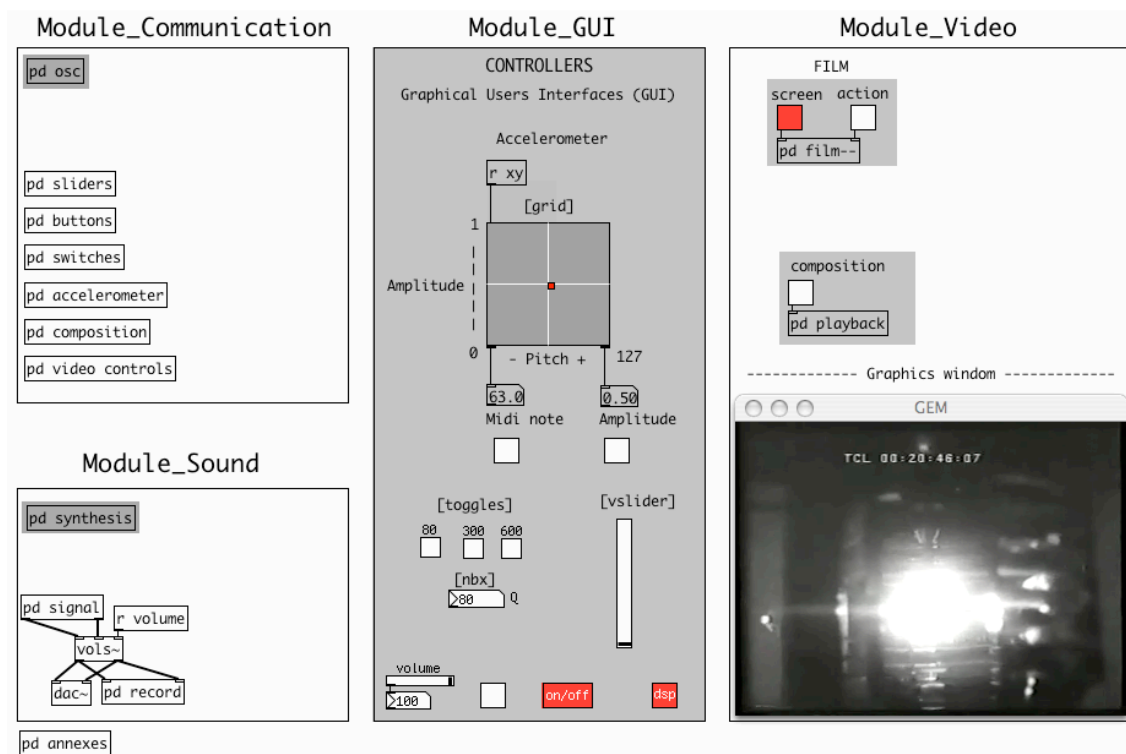


Figura 84. Interface de utilizador (MuDI versão 1.0). Em termos gerais, o Módulo Comunicação recebe, encaminha e trata a informação transmitida pela aplicação do dispositivo electrónico móvel. O Módulo Som está dedicado à geração sonora e à gravação áudio. O Módulo GUI permite visualizar a calibração do sistema, os pontos de referência e a relação entre os gestos realizados e o som gerado. O Módulo Vídeo permite o visionamento do filme e a reprodução musical.

X.2.2.5 Módulo Comunicação

No Módulo Comunicação o *subpatch* “*pd osc*” realiza a recepção de toda a informação transmitida pela aplicação de controlo, enviando-a de seguida para os *subpatches* programados especificamente para receber, encaminhar e tratar individualmente a informação de cada controlador. Estes *subpatches* específicos são: “*pd sliders*”, “*pd buttons*”, “*pd switches*”, “*pd accelerometer*”, “*pd composition*” e “*pd video controls*”.

Neste módulo, os principais *subpatches* são aqueles que se encontram directamente relacionados com o controlo dos principais atributos sonoros usados em contexto musical (i.e., percepção de altura sonora, percepção de intensidade sonora, Timbre, Dinâmicas e Durações).

Dois destes principais *subpatches* são: “*pd sliders*” e “*pd buttons*”. Ambos enviam valores numéricos para controlar o factor Q do filtro, que se encontra no *subpatch* “*pd synthesis*” (Módulo Som). O factor Q (relação entre a frequência central e a largura de banda do filtro) determina a quantidade de ruído filtrado produzindo modificações no espectro sonoro e consequentemente alterações em termos de Timbre.

No *subpatch* “*pd accelerometer*”, os valores das coordenadas do acelerómetro, para movimentos de flexão e extensão do pulso no eixo *x* entre 0° e -180° e de movimentos de rotação interna e externa do pulso, no eixo *y* entre -90° e 90° são escalonados de forma linear, de maneira a se obterem valores entre zero e um (*x-values*) para a amplitude do som e valores entre zero e 127 (*y-values*) para a altura sonora.

Posteriormente, os *x,y-values* são enviados separadamente para o *subpatch* “*pd synthesis*” com o propósito de controlarem a amplitude e a altura sonora do gerador de som. Consequentemente, vão determinar a percepção da intensidade e da altura sonoras.

Os *Switches* A e H (Figura 83) relacionados com o *subpatch* “*pd switches*” permitem manipular a Dinâmica do som de duas maneiras: quando ligamos o *switch* A, o fluxo dos valores do eixo *x* do *subpatch* “*pd accelerometer*” é interrompido e a amplitude do som “congela”. Nesta situação, só podemos manipular a percepção de altura sonora, ou seja, tocar diferentes alturas sonoras com a mesma intensidade (e.g., muito suave, suave, moderado, forte, muito forte). Quando ligamos o *switch* H, o fluxo do eixo *y* é interrompido e a altura sonora “congela”, ou seja, tocamos a mesma altura sonora a diferentes níveis de intensidade sonora (e.g., desde muito suave até muito forte).

Também podemos usar o *switch* H para manipular a Duração do som. Basta para isso pressioná-lo de forma intermitente.

X.2.2.6 Módulo Som

O Módulo Som está exclusivamente dedicado à geração de som e à gravação áudio da composição musical directamente no disco rígido do computador. É representado pelos *subpatches* “*pd synthesis*” e “*pd record*”.

O *subpatch* “*pd synthesis*” é constituído por dois geradores de som. Referimo-nos concretamente a um gerador de ruído branco processado por um módulo de filtragem e de reverberação e pelo Chromasom (Capítulo IX, desta dissertação).

O segundo *subpatch* designado “*pd record*” está dedicado à gravação áudio da composição musical directamente no disco rígido do computador. O ficheiro áudio gravado (um ficheiro estéreo WAV) é armazenado na mesma pasta onde se encontra o ficheiro *Pd-extended* que constitui a aplicação da interface de utilizador, normalmente designado por “*MuDI.pd*”.

X.2.2.7 Módulo GUI

Este módulo permite visualizar a calibração do sistema, os pontos de referência e a relação entre os gestos realizados e o som gerado através dos objectos *Pd-extended* designados por *Graphical Users Interfaces* (GUI). Referimo-nos concretamente aos seguintes objectos: “*grid*”, “*toggle*”, “*vslider*”, “*hslider*” and “*nbx*” (Módulo GUI, Figura 84).

Neste módulo, cada controlador da aplicação do dispositivo electrónico móvel possuiu um objecto GUI específico associado a ele. Quando um controlador é manipulado, o objecto GUI associado reage simultaneamente permitindo, desta maneira, ao intérprete e à audiência acompanhar a relação entre os gestos realizados e o som gerado.

X.2.2.8 Módulo Vídeo

O Módulo Vídeo está dedicado ao visionamento das imagens do filme e à reprodução da composição musical em sincronização com o filme projectado.

Quando o controlador *switch screen* é activado cria-se uma janela gráfica através do objecto “*gemwin*” que se encontra na livraria GEM. A janela gráfica criada permite o visionamento do filme seleccionado.

Tocar e parar o filme são os únicos comandos utilizados para controlar as imagens do filme.

A Figura 85, mostra a relação entre os diferentes módulos da interface de utilizador, os seus elementos, a principal função atribuída a cada um deles e o resultado final em termos de percepção audiovisual.

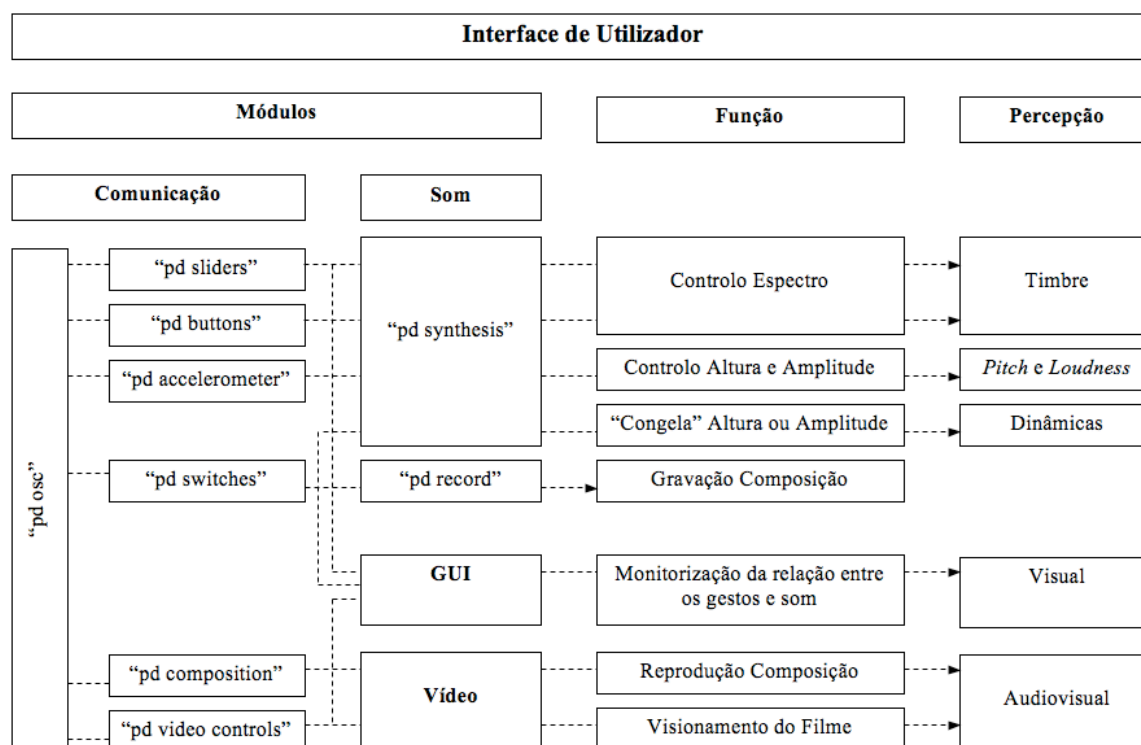


Figura 85. Este diagrama de blocos apresenta (1) a estrutura geral da interface de utilizador, (2) as funções de cada um dos seus elementos e por último (3) o resultado final alcançado em termos de percepção audiovisual. No Módulo Comunicação os *subpatches* "pd sliders", "pd buttons" e "pd accelerometer" estão inteiramente dedicados ao controlo dos geradores de som (Módulo Som). O *subpatch* "pd switches" é multi-funcional, no sentido, em que controla não só a geração do som, mas também a gravação da composição musical. Os *subpatches* "pd composition" e "pd video controls" estão unicamente destinados à reprodução da composição musical em sincronização com o filme seleccionado e com a paragem e reprodução das imagens do filme. Todos os *subpatches*, sem excepção estão ligados ao Módulo GUI para permitirem visualizar a manipulação dos controladores da aplicação do dispositivo electrónico móvel.

X.2.3 Manuseamento

Tocar e deslizar os dedos, flexão, extensão e rotação interna e externa dos pulsos são os principais gestos físicos utilizados para operar o MuDI.

O diagrama de blocos representado na Figura 86, resume a relação entre os controladores, os gestos físicos de interacção, a função atribuída a cada um dos controladores e o resultado final obtido em termos de percepção audiovisual.

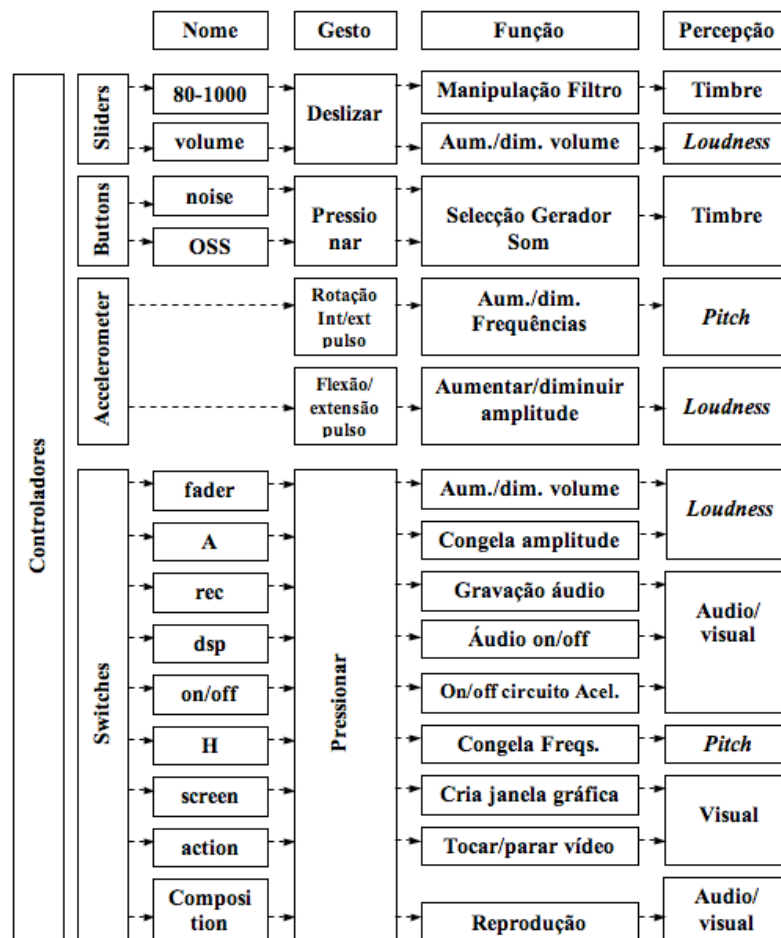


Figura 86. O diagrama de blocos fornece uma visão geral sobre a estrutura e o modo de operação do MuDI (versão 1.0). Através dele podemos ver o nome e o número total de controladores usados para tocar o MuDI. Foram escolhidos dois *sliders*, dois *buttons*, um acelerómetro e nove *switches*. Podemos ainda ver o tipo de gestos físicos usados para os manipular, a função atribuída a cada um deles e o resultado final obtido em termos de percepção audiovisual.

X.2.4 Calibração do MuDI

No modo de execução da aplicação do dispositivo electrónico móvel, o acelerómetro capta a posição do respectivo dispositivo, a velocidade e a aceleração de movimento e emite valores numéricos correspondentes aos ângulos de flexão e extensão dos pulsos do utilizador entre 0° e -180° no eixo x, e valores numéricos correspondentes aos ângulos de rotação interna e externa dos pulsos entre -90° e 90° no eixo y.

Estes valores foram previamente uniformizados através de um processo de tratamento individual da informação enviada pelo acelerómetro, realizado no interior do *subpatch* “*pd accelerometer*”, que se encontra no Módulo Comunicação. Este processo

consistiu no seguinte: dentro dos limites dos ângulos mencionados anteriormente constatou-se que o acelerómetro enviava valores negativos para os respectivos eixos. Para ultrapassar este fenómeno adicionou-se o factor um aos valores recebidos e consequentemente obteve-se um intervalo de números positivos compreendidos entre zero e dois. De seguida, reduziu-se para metade o valor do intervalo mencionado dividindo-o pelo factor dois. Obtiveram-se automaticamente valores lineares compreendidos entre zero e um para a amplitude do som (*x-values*). Por outro lado, ao multiplicarmos os valores do intervalo compreendido entre zero e um, pelo factor 127, obtiveram-se as notas MIDI (*y-values*) que representam a altura sonora.

No entanto, para ajustar os *x,y-values* aos limites dos eixos do objecto “*grid*” (Módulo GUI, Figura 84) foi necessário realizar alguns reajustamentos. O primeiro intervalo de números compreendido entre zero e dois, obtido no início deste processo, para os *x-values* foi dividido pelo factor 1.9 e não pelo factor dois inicial. Em relação aos *y-values*, o intervalo compreendido entre zero e um foi multiplicado, não pelo factor 127, mas por 125.

Desta maneira, ao colocarmos o dispositivo electrónico móvel estático sobre uma mesa, obtém-se um ponto de referência fiável do acelerómetro. Esse ponto de referência representa a calibração do sistema do MuDI e corresponde ao valor 0.5 para a amplitude e à nota MIDI 63 para a altura sonora (Figura 87).

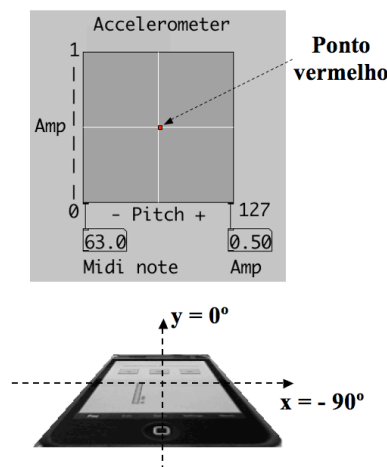


Figura 87. Representação do ponto de referência fiável do acelerómetro (ponto vermelho) e do dispositivo electrónico móvel colocado sobre uma mesa. A posição central do ponto de referência representa a calibração do sistema do MuDI. As suas coordenadas são 0° no eixo y e -90° no eixo x . Em termos de informática musical, as coordenadas correspondem a 0.5 de amplitude e à nota MIDI 63 de altura sonora.

Para uma melhor compreensão do processo de calibração anteriormente descrito apresentamos a Figura 88.

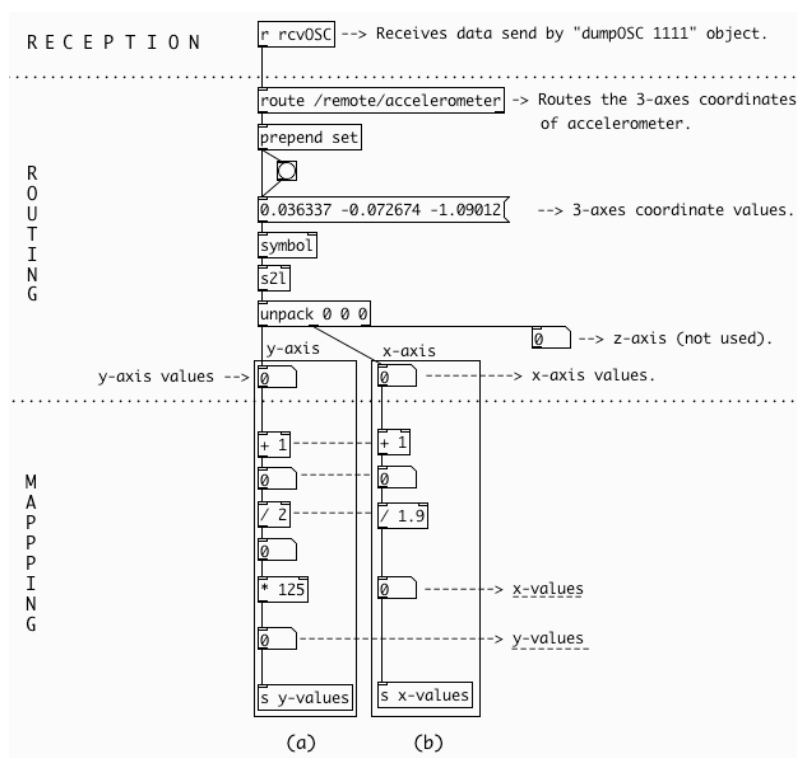


Figura 88. Representação dos processos de recepção, encaminhamento e tratamento dos valores dos eixos x e y transmitidos pelo acelerômetro realizados no interior do subpatch "pd accelerometer" (Módulo Comunicação). A nota (a) representa as operações matemáticas e o processo de reajustamento realizados para obter os y -values (0 – 127, notas MIDI). A nota (b) representa as operações matemáticas e o processo de reajustamento realizados para se obterem os x -values (0 – 1, amplitude).

Os x, y -values obtidos são enviados para o subpatch "pd synthesis" para controlarem a percepção da intensidade e da altura sonoras, e ainda para o objecto "grid" (Módulo GUI), para serem monitorizados.

Foram seleccionados três pontos de referência para cada um dos eixos. Para o eixo y ($y1$, $y2$ e $y3$) e para o eixo x ($x1$, $x2$ e $x3$). A próxima sequência de imagens mostra os respectivos pontos de referência seleccionados para cada um dos eixos (Figura 89 e Figura 90).

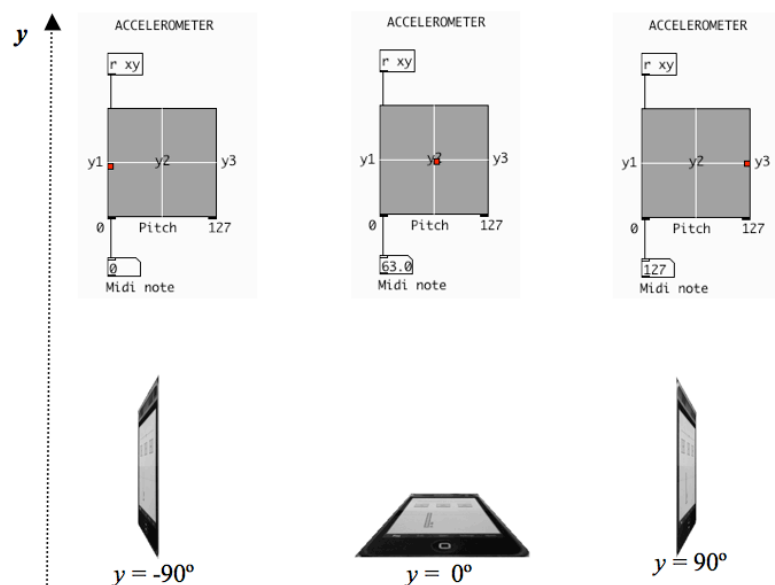


Figura 89. Representação dos pontos de referência (ponto vermelho) para a altura sonora e respectivos ângulos de rotação interna e externa do pulso realizados ao longo do eixo y .

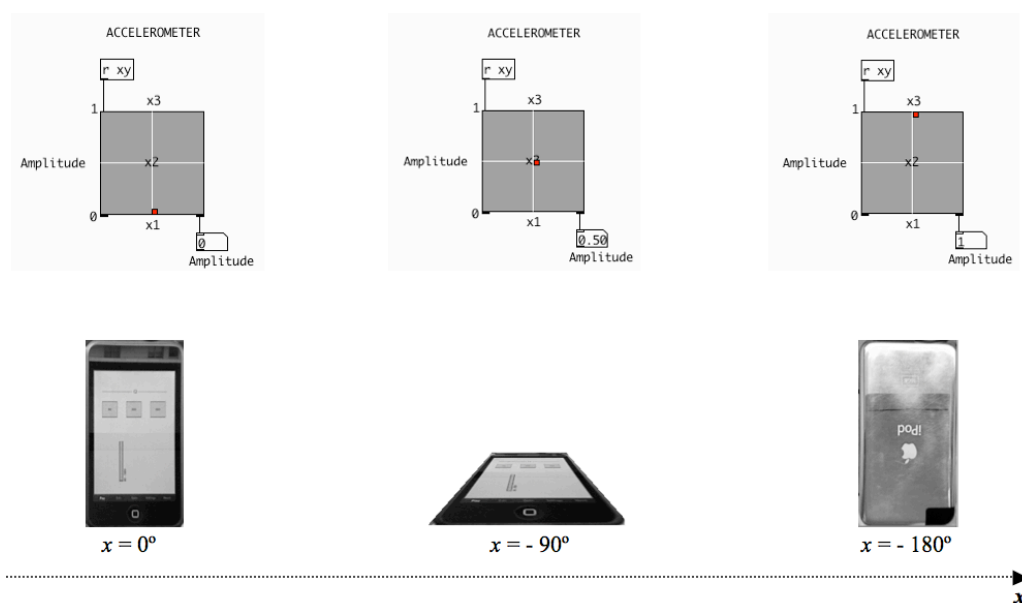


Figura 90. Representação dos pontos de referência (ponto vermelho) para a intensidade sonora e dos respectivos ângulos de flexão e extensão do pulso realizados ao longo do eixo x .

A Tabela 10 e a Tabela 11, relacionam os três pontos de referência obtidos (para a altura sonora no eixo y e para a intensidade sonora no eixo x) e as coordenadas dos movimentos dos pulsos.

Tabela 10. Pontos de referência do eixo y, coordenadas dos movimentos de rotação interna e externa do pulso e nota MIDI correspondente obtida para a altura sonora.

Pontos de referência do eixo y	Coordenadas	Alturas (Número da nota MIDI)
y1	- 90°	0
y2	0°	63
y3	90°	127

Tabela 11. Pontos de referência do eixo x, coordenadas dos movimentos de flexão e extensão do pulso e o valor correspondente obtido para a intensidade sonora.

Pontos de referência do eixo x	Coordenadas	Amplitudes
x1	0°	0
x2	-90°	0.5
x3	-180°	1

Os pontos de referência obtidos fornecem pistas audiovisuais ao compositor/intérprete e poderão ser úteis para tocar o MuDI.

X.2.5 Partitura Vídeo

A possibilidade de gravar directamente no computador o funcionamento da interface de utilizador, através de um programa de captação vídeo de ecrã (e.g., Screenium) permite ao compositor/intérprete obter uma partitura vídeo.

Como qualquer partitura musical, a partitura vídeo obtida funciona como um guia gráfico ou auxiliar de memória. Consequentemente, vai possibilitar reproduzir com bastante precisão os gestos musicais previamente compostos e executados.

Na próxima secção vamos abordar a aplicação do Chromasom como gerador de som em composição musical. Apresentamos dois ciclos de obras compostos integralmente durante o período em que decorreu a nossa investigação. Estes ciclos apresentam as seguintes características comuns: (1) o Chromasom desempenha um papel preponderante na geração de som, (2) o *Pure Data* é usado como ambiente exclusivo de programação, no que respeita, às aplicações das interfaces de utilizador, à síntese de som e à sequenciação musical e (3) o computador portátil é utilizado como auxiliar de composição e instrumento musical.

Na descrição das obras musicais apresentaremos as aplicações concebidas, a sua estrutura, funcionamento, operacionalização e interfaces de utilizador. Mencionaremos ainda, questões conceptuais relacionadas com as escolhas dos títulos e dos subtítulos das obras, e as estratégias de composição adoptadas.

X.3 Ciclo *Música Digital Contemporânea Erudita* (MuDiCE)

Neste ciclo de obras, as aplicações e suas interfaces de utilizador partilham funcionalidades comuns. Todas possuem módulos de controlo, que enviam mensagens de sequenciação, espacialização e mensagens que determinam a percepção de tempo, de ritmo, de dinâmicas, de intensidade e de altura sonoras; módulos áudio, que englobam as fontes sonoras, o processamento de som (reverberação, *delay* e filtragem), os volumes individuais e o volume geral, e ainda, um módulo de gravação, que permite gravar as interpretações, directamente no disco rígido do computador, obtendo-se desta maneira, um ficheiro estéreo WAV.

Em termos de expressão musical procurámos criar gestos musicais, previamente concebidos e posteriormente interpretados. O gesto musical pressupõe uma intenção e uma acção. Uma intenção estético-sonora global, de concepção musical, da responsabilidade do compositor, e uma acção, um gesto físico, do intérprete na interacção com o instrumento musical durante a interpretação da obra.

Quer na perspectiva de compositor, quer na perspectiva de intérprete, consideramos o gesto musical como sendo o percurso, ou o ciclo de vida do som. Nesse percurso, o som nasce algures, desenvolve-se e acaba por se extinguir. É como se do nosso ciclo de vida se tratasse. O exemplo de um gesto musical pode ser encontrado e escutado, no *second moment – the periclitanti*, da obra *perpetual melody – contrasting moments*. No entanto, o gesto musical só se concretiza, quando o ouvinte percebe, ou reconhece na obra que escuta, a intencionalidade do compositor. Esta encontra-se musicalmente materializada na interpretação da obra. O gesto musical é aplicável a qualquer parâmetro sonoro ou musical. Para além disso, uma obra musical pode incorporar vários gestos, um único gesto, ou ainda, pressupor, se essa for a intenção do compositor, não ter gesto nenhum.

Em termos de etapas metodológicas começamos por (1) testar combinações timbricas entre o Chromasom e outras fontes sonoras, (2) restringir as fontes sonoras ao

Chromasom, (3) gerar massas sonoras e (4) utilizar os timbres resultantes da aplicação do Chromasom a um circuito básico de síntese por FM.

Neste ciclo destacamos as seguintes obras musicais:

- *mudice\$1 – música de uma tecla só (a pretensiosa)*, por estrear;⁵⁷
- *mudice\$2 – a irreversível*, estreada no concerto de comemoração dos 25 anos da Miso Music Portugal, em Abril de 2010;⁵⁸
- *mudice\$3 – a perpétua*, estreada no Festival Música Viva, em Setembro de 2010;⁵⁹
- *Perpetual Melody – contrasting moments*, por estrear.⁶⁰

X.3.1 mudice\$1 – música de uma tecla só (a pretensiosa).

Mudice\$1 – música de uma tecla só (a pretensiosa) resultou de uma encomenda feita pela Miso Music Portugal, para ser integrada numa obra colectiva intitulada *Cadavre Exquis*. A sua estreia aconteceu no concerto de comemoração dos 25 anos da instituição mencionada.

Nesta obra utilizamos somente uma tecla do teclado do computador para despoletar o som, razão pela qual, lhe atribuímos o subtítulo de “*música de uma tecla só (a pretensiosa)*”.

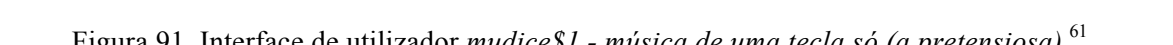
Entre cada gesto musical, a altura sonora e o tempo de execução aumentam de forma progressiva até se atingir o limite de tensão. A partir desse momento, a tensão é libertada, a percepção da altura sonora mantém-se constante e confortável e os gestos musicais voltam novamente a respirar.

⁵⁷ A obra musical pode ser escutada no DVD em anexo, no directório MuDICE, com o nome *mudice\$1 – música de uma tecla só (a pretensiosa).wav*

⁵⁸ A obra musical pode ser escutada no DVD em anexo, no directório MuDICE, com o nome *mudice\$2 – a irreversível.wav*

⁵⁹ A obra musical pode ser escutada no DVD em anexo, no directório MuDICE, com o nome *mudice\$3 – a perpétua.wav*

⁶⁰ A obra musical pode ser escutada no DVD em anexo, no directório MuDICE, com o nome *Perpetual Melody – contrasting moments.wav*



leitura da melodia em ciclo. O *módulo 1c*, é dedicado exclusivamente à transposição da percepção de altura sonora do Chromasom. No *subpatch “pd novos_intervalos”* (*módulo 1c*) os primeiros quatro “bangs” permitem-nos escolher diferentes relações intervalares, enquanto que o “toggle auto” inicia um módulo automático de relações intervalares aleatórias; no *subpatch “pd 3M_6M_7M_8”* seleccionamos relações intervalares por terceiras, sextas e sétimas maiores. Nesta versão da obra, em relação à voz “*pd Chromasom*” funciona por defeito, a relação intervalar de equivalência de oitava. O *módulo 1d*, armazena a melodia numa tabela. O *módulo 1e*, transpõe em termos de percepção de altura sonora a voz “*pd noise*”. O *módulo 1f*, controla a transposição geral da melodia, em termos de percepção de altura sonora e o *módulo 1g*, regula o volume geral. Os módulos áudio *2a*, *2b*, *2c*, *2d* e *2e* são geradores de som. Estes geradores de som vão determinar as vozes musicais. Assim, o *módulo 2a*, *subpatch “pd Chromasom”*, é constituído por um Chromasom. O *módulo 2b*, *subpatch “pd som_puro”* é formado por um oscilador. O *módulo 2c*, *subpatch “pd noise”* é um gerador de ruído branco (*noise~*) constituído por onze componentes separadas por intervalos de terceira maior (Figura 92).

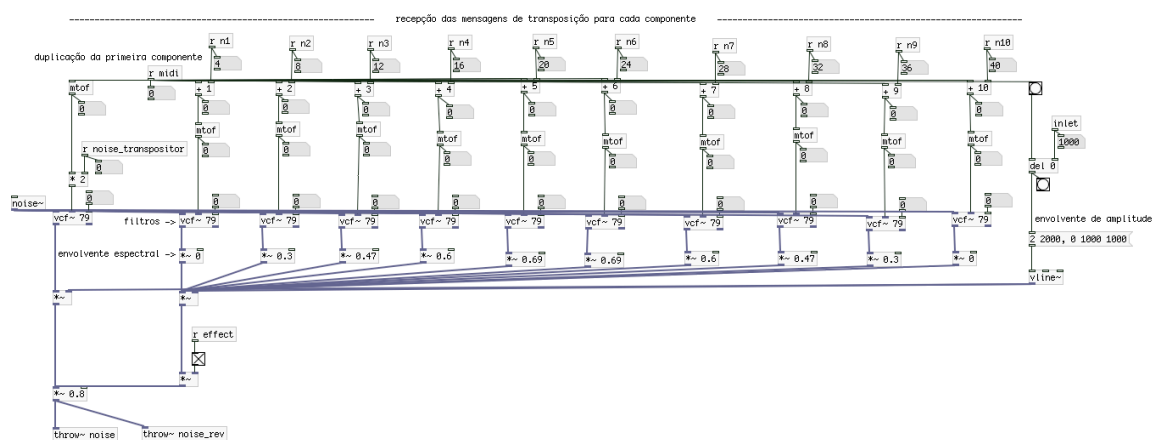


Figura 92. Gerador de som *subpatch “pd noise”*.

Cada uma das componentes da voz “*pd noise*” é processada por um filtro Passa-Banda. A frequência central de cada filtro determina a percepção de altura sonora. Foi-lhe aplicada uma envolvente espectral simétrica determinada de forma logarítmica. A primeira componente está duplicada podendo ser transposta em termos de percepção de altura sonora, através da manipulação do *módulo transposição “pd noise”* (Figura 91,

módulo 1e)⁶². O *módulo 2d*, *subpatch* "pd osciladores" é formado por um som complexo constituído pelos cinco primeiros harmónicos da série harmónica completa (Figura 93).

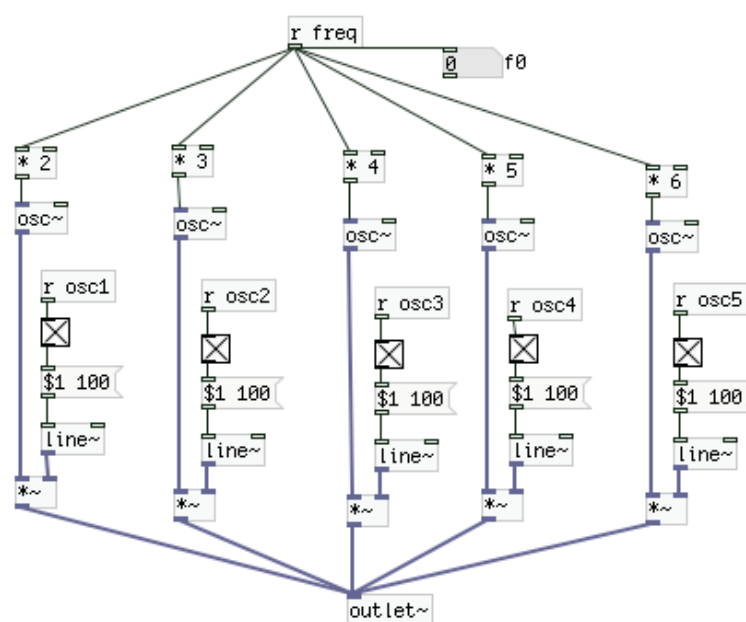


Figura 93. Gerador de som *subpatch* "pd osciladores".

O *módulo 2e*, *subpatch* "pd guit" representa o som de uma guitarra acústica sintetizada a partir do algoritmo Karplus-Strong.⁶³

O *módulo 2f*, recebe o sinal áudio e determina o volume e a espacialização de cada um dos geradores sonoros, atrás descritos. O *módulo 2g*, determina o volume e a quantidade de reverberação dos mesmos geradores sonoros. O *módulo 3*, permite a gravação da interpretação num ficheiro áudio estéreo WAV, ficheiro esse, que fica armazenado no disco rígido do computador.

O diagrama de blocos representado na Figura 94, oferece-nos uma visão geral sobre a organização dos módulos e a maneira como eles comunicam uns com os outros, quer em termos de mensagens de controlo, quer de encaminhamento áudio.

⁶² Refira-se, a título de curiosidade que, apesar do *módulo 1e*, ter sido explorado e de ter produzido resultados musicais interessantes não foi usado nesta versão. A exigência da duração obrigatória da encomenda (três minutos), levou-nos a adoptar como estratégia de composição reduzir, não só os recursos, como também o discurso musical ao seu essencial.

⁶³ <http://www.lubbertdas.org/pd/>

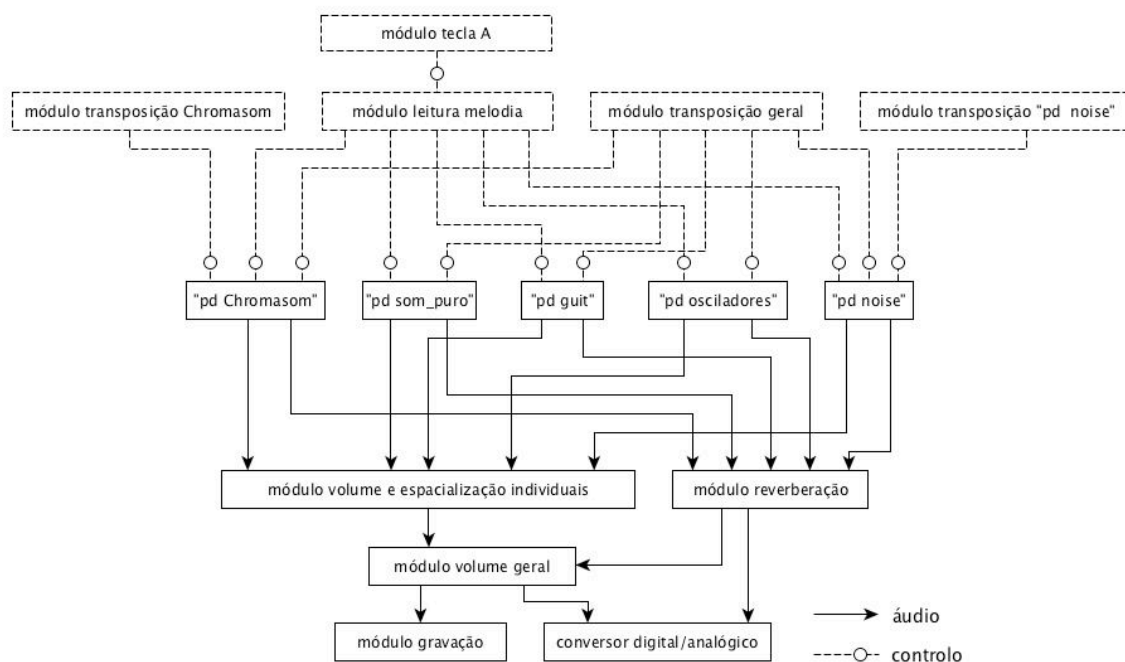


Figura 94. Diagrama de blocos da interface de utilizador *mudice\$1* – música de uma tecla só (a pretensiosa).

De acordo, com o diagrama de blocos apresentado na Figura 94, o *módulo tecla "A"*, envia mensagens para o módulo de leitura da melodia. Este envia notas MIDI para cada uma das fontes sonoras. O *módulo transposição geral* envia mensagens de transposição para todas as fontes sonoras. O *módulo transposição Chromasom* está dedicado exclusivamente ao Chromasom e o *módulo transposição "pd noise"* à voz "*pd noise*". Cada gerador de som recebe as notas MIDI da melodia e as mensagens de transposição. O áudio produzido pelas fontes sonoras é encaminhado simultaneamente para o *módulo volume e espacialização individuais* e para o *módulo reverberação*. O áudio proveniente do *módulo volume e espacialização individuais* é dirigido para o *módulo volume geral* e este, por sua vez, vai ao mesmo tempo para o *módulo gravação* e para o *conversor digital-analógico*. O áudio processado pelo *módulo reverberação* é enviado simultaneamente para o *módulo volume geral* e para o *conversor digital-analógico*.

Em relação à operacionalização, após se lançar a interface de utilizador, o intérprete, para despoletar o som deverá premir a tecla "A", do teclado do computador. Esta tecla foi escolhida por permitir o apoio do pulso e da palma da mão sobre a superfície que se encontra do lado esquerdo do computador, permitindo dessa maneira, o conforto e a estabilidade necessárias à interpretação. Por outro lado, a mão direita, é

utilizada para operar o *módulo transposição geral* aplicado à percepção da altura sonora (*módulo 1f*). Nesta versão, durante a gravação da obra, interagimos somente com o *módulo gravação*, com a tecla “A”, e com o *módulo transposição geral*. Até sensivelmente a meio da obra, entre cada gesto musical, realizamos com a mão esquerda alterações rítmicas e de tempo e simultaneamente, com a mão direita transposições em termos de percepção de altura sonora, premindo para isso, da esquerda para a direita, cada uma das mensagens do *módulo transposição geral* (*módulo 1f*). A partir do meio da obra, seleccionamos a mensagem de transposição “72”, mantendo, desta maneira, a percepção de altura sonora constante até ao final.

A tecla “A”, ao ser premida, faz disparar um “bang”. Este por sua vez, faz avançar, por ordem numérica crescente e em ciclo, valores compreendidos entre zero e 35. Está associado a cada valor numérico uma nota MIDI da melodia (Figura 95).



Figura 95. Melodia da *mudice\$1 – música de uma tecla só (a pretensiosa)*.

A melodia está armazenada numa tabela. A nota MIDI é convertida em frequência e enviada para cada um dos geradores de som, de maneira a determinar a percepção de altura sonora. Após se libertar a tecla “A”, surge automaticamente, com um atraso de 1000 milissegundos, a voz “*pd noise*” com a duração máxima de 4000 milissegundos. A duração da voz “*pd noise*” é interrompida, sempre que se prime, novamente a tecla “A”. À excepção da voz “*pd noise*”, todas as outras vozes são tocadas em uníssono. Este procedimento permitiu-nos combinar e confrontar o comportamento tímbrico do Chromasom com os timbres das outras vozes, no sentido, de constataremos, se globalmente o resultado tímbrico era interessante.

Por razões de contexto musical e de um acontecimento familiar irreversível decidimos estrear no concerto de comemoração dos 25 anos da Miso Music Portugal em Abril de 2010, não a *mudice\$1 – música de uma tecla só (a pretensiosa)*, mas uma versão intitulada *mudice\$2 – a irreversível*.

X.3.2 *mudice\$2* – a irreversível.

Nesta obra, numa primeira etapa, processámos em tempo real, através de um módulo granular, designado por “*dlgranulator*”⁶⁴ o ficheiro áudio estéreo WAV da *mudice\$1* – *música de uma tecla só (a pretensiosa)*, que representa a primeira voz. Desse processamento, resultou a segunda voz (som granulado). Em termos estéticos, a primeira voz representa a linha da vida e a segunda, a linha da morte. Na sequenciação áudio, colocámos ambas as vozes em paralelo. No entanto, entre a primeira e a segunda voz existe um atraso de seis milissegundos, produzindo-se desta maneira, um *Canon* a duas vozes. Aplicámos às dinâmicas das vozes as envolventes que se encontram representadas na Figura 96.

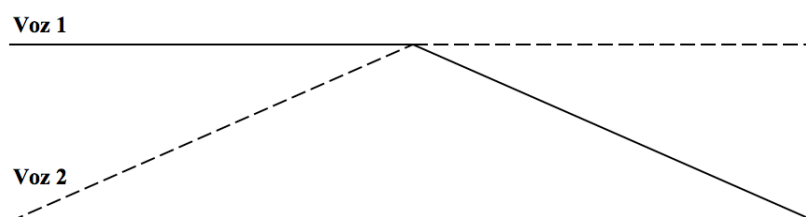


Figura 96. Representação gráfica das envolventes das dinâmicas aplicadas às vozes da *mudice\$2* – *a irreversível*. A linha contínua corresponde à envolvente aplicada à primeira voz (linha da vida), enquanto que a linha a tracejado diz respeito à envolvente aplicada à segunda voz (linha da morte). O conjunto das envolventes adquire a imagem de uma mesa de perfil, ou de um K deitado facialmente.

Metaforicamente, descrevemos as dinâmicas representadas na Figura 96, da seguinte maneira: ambas as vozes começam ao mesmo tempo, porque só há morte enquanto há vida. Na corrida de sentido único, que é a nossa existência, há um momento em que a segunda voz se sobrepõe à primeira. A partir desse instante, sem possibilidade de retorno, a primeira voz vai progressivamente diminuindo de intensidade, enquanto que

⁶⁴ <http://www.pdpatchrepo.info/hurlleur/dlgranulator.zip>

a segunda se mantém constante e determinada a chegar ao fim. A morte é o paradigma do irreversível.

A interface de utilizador da obra está representada na Figura 97.

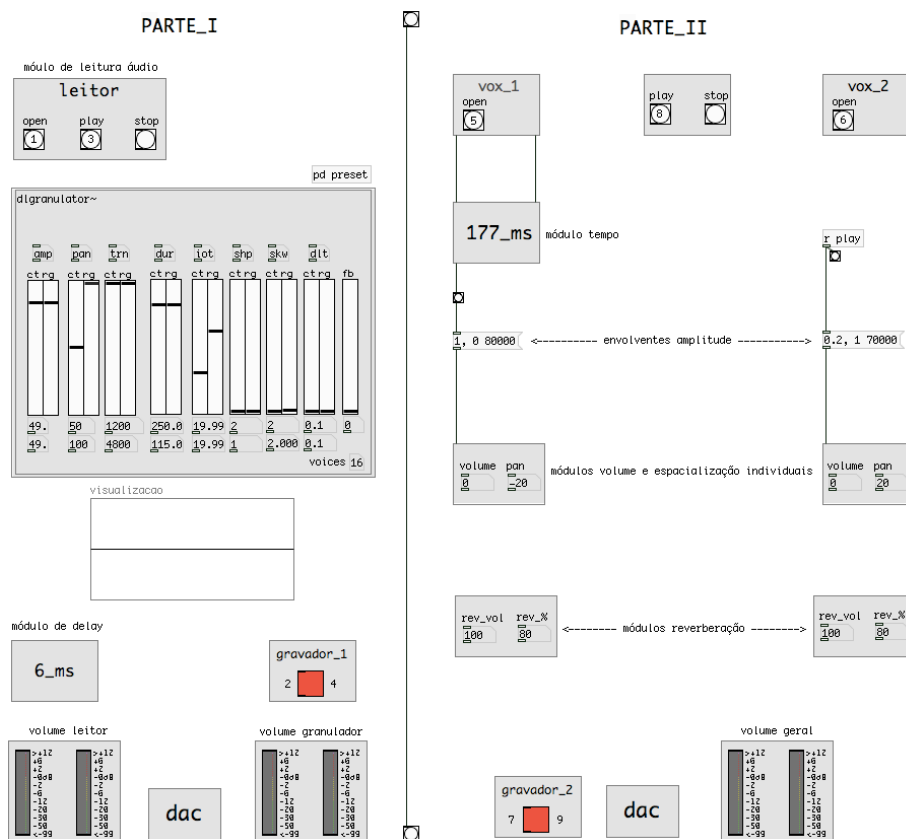


Figura 97. Interface de utilizador *mudice\$2 – a irreversível*.

A Parte I, é constituída pelos módulos: *leitura áudio*, *granulador* (dlgranulator~), *delay*, *gravador_1*, *volume leitor*, *volume granulador* e *conversor digital-analógico*. A Parte II, é constituída pelos módulos: *vox_1*, *vox_2*, *play/stop*, *tempo*, *volume e espacialização individuais*, *reverberação*, *volume geral*, *gravador_2* e *conversor digital-analógico*.

O diagrama de blocos representado na Figura 98, permite-nos observar como os módulos se encontram organizados e comunicam entre si, em termos de mensagens de controlo e encaminhamento áudio.

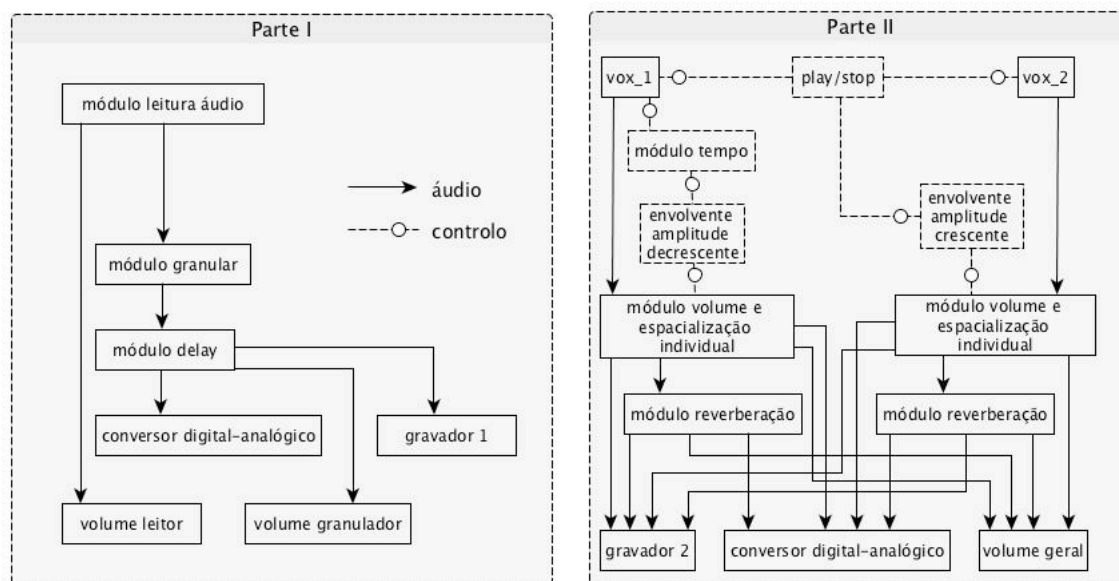


Figura 98. Diagrama de blocos da interface de utilizador *mudice\$2 – a irreversível*.

A Parte I, está totalmente dedicada à produção áudio. O áudio do *módulo leitura* é encaminhado simultaneamente para o *módulo volume leitor* e para o *módulo granular*. O áudio que resulta do processamento do *módulo granular* (som granulado) é enviado para o *módulo delay*. Este módulo produz um atraso de 6000 milissegundos no envio do áudio para os *módulos volume granulador*, *conversor digital-analógico* e *gravador_1*.

Na Parte II, em ambos os leitores (*vox_1* e *vox_2*), o áudio é encaminhado para os respectivos *módulos volume e espacialização individuais*. Destes módulos é simultaneamente enviado para os *módulos gravação_2*, *conversor digital-analógico*, *volume geral* e *reverberação*. O áudio proveniente dos *módulos de reverberação* é também enviado para os *módulos de gravação_2*, *conversor digital-analógico* e *volume geral*.

São os módulos de controlo que determinam as envolventes de amplitude aplicadas às vozes que se encontram representadas na Figura 96. Assim, o *módulo tempo* controla a envolvente de amplitude decrescente aplicada à primeira voz, enquanto que o *módulo play/stop* controla a envolvente de amplitude crescente aplicada à segunda voz. A amplitude linear da primeira voz começa com o valor um e aos 70000 milissegundos, decresce progressivamente, durante 80000 milissegundos até zero, enquanto, que a amplitude linear da segunda voz começa com o valor zero ponto dois e aumenta progressivamente durante 70000 milissegundos até ao final da obra.

Na Parte I, a operacionalização do *módulo granular* foi realizada em tempo real, a partir dos valores iniciais dos parâmetros do granulador, que se encontram pré-definidos e gravados no *subpatch* “*pd presets*”, que se encontra no canto superior direito, por cima do respectivo granulador (Figura 97). Consequentemente, o resultado sonoro gravado no disco rígido do computador é único e irreversível, no sentido, de não ser possível corrigir as opções musicais tomadas nem os valores dos parâmetros do *módulo granular*, que foram sendo manipulados durante a interpretação. Esta, é também uma das razões, pela qual atribuímos à obra o subtítulo de “*irreversível*”.

Na Parte II, começamos por abrir o ficheiro áudio *mudice\$1 – música de uma tecla só (a pretensiosa)*, através do leitor “*vox_1*” e o ficheiro áudio que contém o som granulado, através do leitor “*vox_2*”. Depois, ligamos o *módulo gravação* e pressionamos a mensagem “*play*” (*módulo play/stop*) para darmos início à obra. Até ao final, não há mais nenhuma intervenção por parte do intérprete, limitando-se este, após o som geral se extinguir, a desactivar o *módulo gravação_2*.

X.3.3 mudice\$3 – a perpétua.

Mudice\$3 – a perpétua foi seleccionada para integrar a instalação sonora colectiva *Sound Walk*, no Festival Música Viva de 2010. De acordo, com o regulamento do festival, deveria ter a duração compreendida entre os dois e os cinco minutos.

O subtítulo desta obra deriva, do facto, de utilizarmos como conteúdo musical a melodia que constitui a Ilusão Sonora da Melodia Perpétua (Figura 77). Ao contrário das duas obras anteriores restringimos as fontes sonoras unicamente ao Chromasom.

Com esta obra musical pretendemos testar a capacidade do Chromasom em gerar texturas sonoras densas. Para isso, utilizámos onze Chromasons, um por cada nota musical da escala dividida em 12 partes iguais, ou seja, um Chromasom para o C, outro para o C#, e assim sucessivamente até B. As texturas sonoras densas foram conseguidas através da utilização simultânea dos vários Chromasons, da leitura da melodia (Figura 77), num tempo muito rápido e através da aplicação de valores máximos ao parâmetro *liveness* do módulo de reverberação. Na textura densa da obra é perceptível, a rugosidade ou a aspereza sonoras, originadas pelos batimentos das componentes dos sons que se encontrarem próximas em termos de frequências.

A interface de utilizador usada durante a interpretação encontra-se representada na Figura 99.

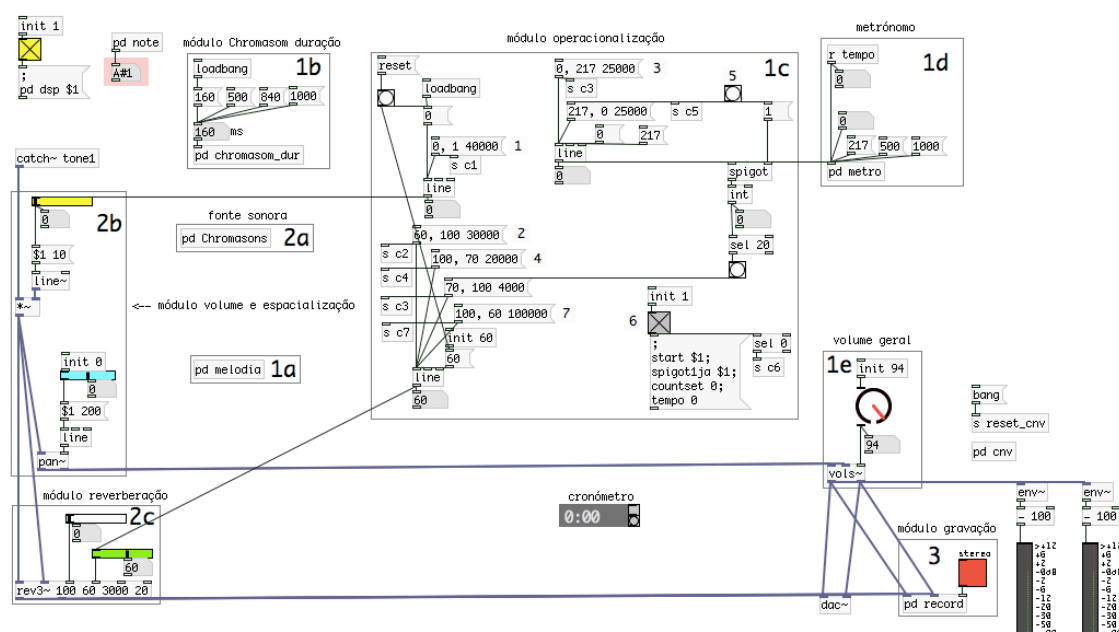


Figura 99. Interface de utilizador *mudice\$3 - a irreversible*.⁶⁵ A classificação número um, representa os módulos de controlo; a número dois, os módulos áudio e a número três, o módulo de gravação.

Em relação aos módulos de controlo, no *módulo 1a* a melodia encontra-se armazenada numa tabela. O *módulo 1b* determina a duração de cada Chromasom. O *módulo 1c* é o elemento de controlo mais importante, porque permite interpretar os gestos musicais concebidos e consequentemente obter a expressão musical desejada. Quanto ao *módulo 1d*, este é um metrónomo que controla a velocidade de leitura da melodia e, finalmente, o *módulo 1e* que controla o volume geral. No que concerne aos módulos áudio, o *módulo 2a* representa a fonte sonora; o *módulo 2b* determina o volume e a espacialização individuais; o *módulo 2c* está dedicado à reverberação e o *módulo 3*, à gravação.

O diagrama de blocos representado na Figura 100, permite-nos obter uma imagem geral da organização dos módulos e da sua comunicação, em termos de mensagens de controlo e de encaminhamento áudio.

⁶⁵ A operacionalização da interface de utilizador pode ser visualizada no DVD em anexo, no directório MuDICE, com o nome *mudice\$3 - a perpétua_interface.mov*

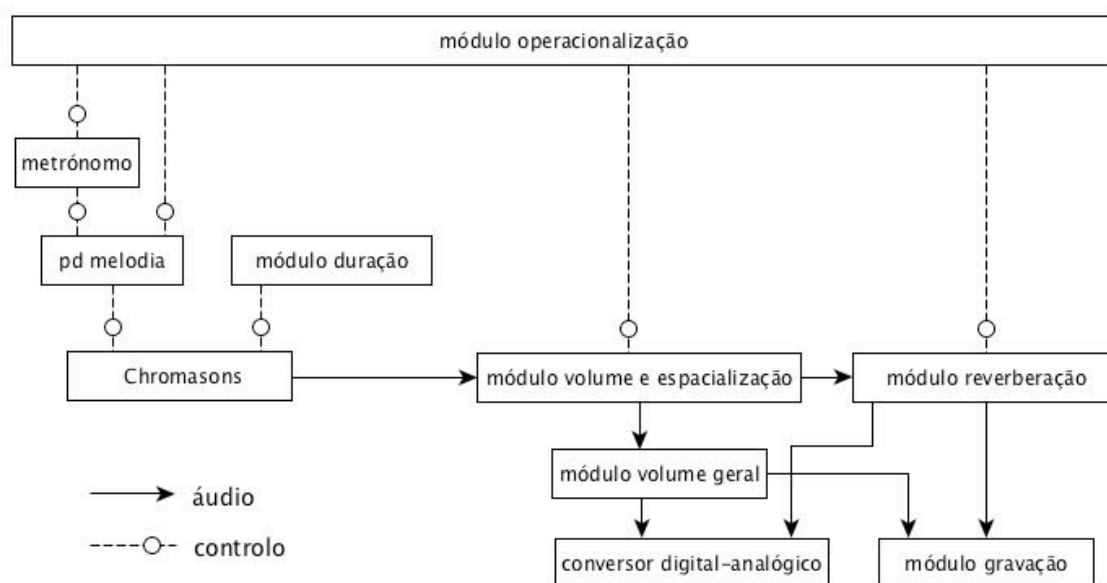


Figura 100. Diagrama de blocos da interface de utilizador *mudice\$3 – a perpétua*.

De acordo, com o diagrama de blocos apresentado na Figura 100, o *módulo operacionalização* envia mensagens para o metrónomo de maneira a determinar a velocidade de leitura da melodia. Paralelamente, o *módulo duração* vai determinar a duração sonora dos Chromasons. O áudio produzido pelos Chromasons é encaminhado para o *módulo volume e espacialização individuais*. Deste módulo, vai simultaneamente para o *módulo volume geral* e para o *módulo reverberação*. A partir daqui, quer o áudio do *módulo volume geral*, quer do *módulo reverberação* vão simultaneamente para o *conversor digital-analógico* e para o *módulo gravação*.

Na operacionalização da interface de utilizador, a melodia, que se encontra armazenada numa tabela (*módulo 1a*), é lida automaticamente, limitando-se o intérprete a seguir os passos numerados por ordem crescente, activando e desactivando botões e mensagens que são enviadas às envolventes aplicadas ao tempo e ao parâmetro *liveness* do *módulo reverberação*, tendo a preocupação de avançar para o passo seguinte, quando o valor das mensagens de controlo atingem o valor máximo, ou de acordo, com a sua sensibilidade musical.

Assim, a interacção do intérprete com o *módulo operacionalização* consiste em sete passos: no primeiro passo, é aplicada uma envolvente de amplitude crescente (*fade-in*) ao volume individual da fonte sonora; no segundo passo, elevamos progressivamente o parâmetro *liveness* do *módulo reverberação* até 100%; no terceiro passo, desaceleramos

o tempo de leitura da melodia; no quarto passo, diminuimos progressivamente o parâmetro *liveness* do *módulo reverberação* até 70%; no quinto passo, aceleramos o tempo até valores de velocidade muito rápidos; no sexto passo, desligamos a automatização da leitura da melodia e esperamos alguns segundos; no sétimo passo, aplicamos uma envolvente decrescente (*fade-out*) ao parâmetro *liveness* do *módulo reverberação* e com isso, obtêm-se uma diminuição suave na percepção da intensidade sonora. A partir desse momento e até desligarmos o *módulo gravação*, gesto esse, que simboliza o fim da obra, não intervimos mais e deixamos que o som se extinga.

X.3.4 Perpetual Melody – contrasting moments.

O título *Perpetual Melody – contrasting moments* deriva do facto de se usar como conteúdo musical a melodia que constitui a Ilusão Sonora da Melodia Perpétua (Figura 77) e ainda, porque os sete *moments* (i.e., os sete andamentos) que a constituem, serem contrastantes entre si, em termos de percepção de altura sonora, de tempo, de envolventes de amplitude, etc. Assim no *first moment – the inert*, aplicamos uma envolvente ao tempo, de maneira a criar um desacelerando, desde muito rápido até lento. No *second moment – the periclitanti*, utilizamos um tempo rápido e ataques curtos. No *third moment – the messianic*, experimentamos o oposto, ou seja, utilizamos um tempo lento com ataques e decaimentos longos. No *fourth moment – the second solist*, apresentamos um tempo rápido e um registo sonoro grave. No *fifth moment – the first solist*, temos precisamente o oposto, ou seja, um tempo lento e um registo sonoro sobre-agudo. No *sixth moment – the burlesque*, utilizamos síntese por modulação de frequência, ou seja, modulamos um Chromasom com outro Chromasom. No *seventh moment – the duet*, na tentativa de criar contrastes, não só em termos de alturas sonoras, mas também de tempo, envolventes de amplitude, etc., utilizamos duas vozes, uma voz modulada por FM e outra não modulada.⁶⁶

A obra utiliza como fonte sonora exclusivamente o Chromasom. Este, foi testado até ao seu limite, em termos de percepção de altura sonora, Timbre, Tempo e envolventes de amplitude, para além de ter sido submetido a um circuito básico de síntese por FM com

⁶⁶ A interface de utilizador desta composição, que se encontra representada na Figura 101, pode ser visualizada no DVD em anexo, no directório MuDICE, com o nome *perpetual melody – contrasting moments_interface.mov*

o objectivo de criar novos timbres e de produzir segregação sonora entre as vozes musicais utilizadas (Chowning 1999, pp. 267).

Como gestos musicais foram usadas transposições aplicadas à percepção de altura sonora, envolventes aplicadas à percepção de intensidade sonora, envolventes de tempo e envolventes de amplitude aplicadas especificamente, aos ataques e decaimentos dos sons. Para além disso, aplicámos envolventes aos principais parâmetros usados no circuito básico de síntese por FM, ou seja, à frequência portadora, à frequência moduladora e ao índice de modulação.

Foi criada uma aplicação com uma interface de utilizador para se escutar a respectiva obra musical e consequentemente observar os aspectos composicionais mencionados (Figura 101).



Figura 101. Interface de utilizador *Perpetual Melody – contrasting moments*.⁶⁷

Foram criadas quatro vozes. A primeira e a segunda voz, “vox1” e “vox2” são constituídas, cada uma delas, por onze Chromasoms. A terceira voz, “vox3” é constituída por um único Chromasom e a quarta voz, “vox4” é constituída por um circuito básico de síntese por FM aplicado ao Chromasom.

Assim, no que respeita à distribuição das vozes, no *first moment – the inert*, só é utilizado uma voz, “vox1”. No *second moment – the periclitanti*, são utilizadas duas vozes, “vox1” e “vox2”. No *third moment – the messianic*, são utilizadas duas vozes, “vox1” e “vox4”. No *fourth moment – the second solist* e no *fifth moment – the first solist*, só é utilizado uma voz, “vox3”. No *sixth moment – the burlesque*, também é utilizada uma

⁶⁷ A interface de utilizador pode ser consultada no DVD em anexo, no directório MuDICE, com o nome *perpetual_melody_contrasting_moments.pd*

única voz, “vox4” e no *seventh moment – the duet*, são utilizadas duas vozes, “vox3” e “vox4”.

Cada uma das vozes possuiu um *módulo volume e espacialização*; um *módulo reverberação*; um *módulo envolventes de amplitude*, que controla o ataque e o decaimento dos sons; um *módulo tempo* e um *módulo controlo*. A partir do *módulo controlo*, através do envio automático de mensagens é manipulada a percepção de altura sonora, as envolventes de amplitude, a espacialização, os parâmetros de reverberação (*output level* e *liveness*), os volumes individuais, o volume geral e a duração de cada *moment*.

O diagrama de blocos representado na Figura 102, permite-nos obter uma imagem geral da organização dos módulos e da sua comunicação, em termos de mensagens de controlo e de encaminhamento áudio.

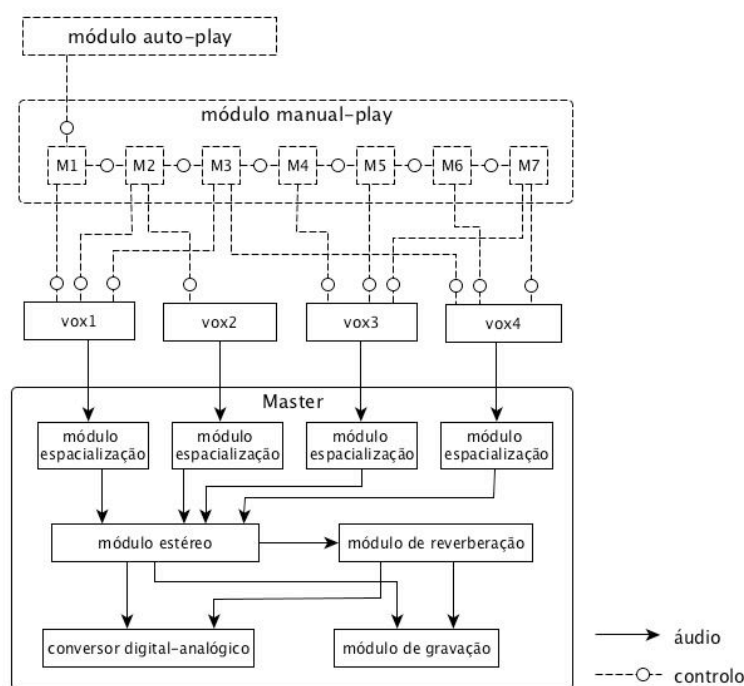


Figura 102. Diagrama de blocos da interface de utilizador *perpetual melody – contrasting moments*.

O áudio produzido por cada uma das vozes é encaminhado para o *módulo Master* para ser, em primeiro lugar, espacializado e depois, reunido em dois canais de saída, de maneira a produzir estereofonia. O áudio geral estéreo é processado pelo *módulo reverberação* e enviado simultaneamente para o *módulo gravação* e para o *conversor digital-analógico*.

Para escutar a obra, a interface de utilizador, através dos módulos de controlo *auto play* e *manual play*, possibilita dois tipos de operacionalização: uma automática, que permite ouvir sequencialmente e por ordem numérica os *moments*, bastando para isso, accionar o botão “*auto_play*” e uma manual, permitindo ao utilizador escutar os *moments*, de acordo, com o seu critério de preferência (Figura 101).

X.4 Ciclo Sonificação – Filmes para Música

No que diz respeito ao contexto da interpretação musical com dispositivos electrónicos móveis e dos novos paradigmas da composição musical em tempo real, através do MuDI iniciámos um novo ciclo de obras designado por “*sonificação – filmes para música*”.

Deste ciclo, fazem parte duas obras. A primeira, chama-se “*sonificação – filmes para música\$1*”,⁶⁸ onde usamos imagens extraídas do filme *Douro, Faina Fluvial* (1931) de Manoel de Oliveira e foi estreada em Setembro de 2009, no 12.º Simpósio Brasileiro de Computação Musical (SBCM).⁶⁹

A segunda, “*sonification – films for music\$2*”⁷⁰, onde utilizamos imagens extraídas do filme *Metropolis* (1927) de Fritz Lang, foi apresentada (versão curta) em Outubro de 2011, no 2.º Fórum Internacional Itinerários Musicais – Música e Gesto, Universidade Nova de Lisboa, Portugal,⁷¹ e em Maio de 2012, na 12.ª Conferência Internacional *New Interfaces for Musical Expression* (NIME), Universidade de Michigan, Ann Arbor – Michigan, EUA.⁷² A descrição, a estrutura modular e a operacionalização das interfaces de utilizador encontram-se no Sub-subcapítulo X.2.2.4 – Interface de Utilizador da Aplicação de Computador, da nossa dissertação.

⁶⁸ A composição pode ser consultada no DVD em anexo, no directório MuDI, com o nome *sonificação – filmes para música\$1.mov*

⁶⁹ Um excerto da estreia pode ser visualizado em <https://www.youtube.com/watch?v=voD2vFEoXJA> ou visualizar o concerto no DVD em anexo, no directório MuDI, com o nome *MuDI_sbcm_2009.avi*

⁷⁰ A composição pode ser consultada no DVD em anexo, no directório MuDI, com o nome *sonification – films for music\$2.mov*

⁷¹ A apresentação pode ser visualizada em <http://www.youtube.com/watch?v=dD4F8FQyPcA> ou no DVD em anexo, no directório MuDI, com o nome *MuDI_itinerários_musicais_2011.mov*

⁷² A apresentação pode ser visualizada em <http://www.youtube.com/watch?v=fAKAsq2A3FE> ou no DVD em anexo, no directório MuDI, com o nome *MuDI_nime_2012.mov*

No processo de composição, as imagens do filme funcionam como pistas musicais, no sentido, em que podem sugerir ao compositor determinados gestos. Por exemplo, quando começar ou acabar um som (Duração), fazer uma pausa ou um silêncio, tocar tempos lentos ou rápidos (Tempo), tocar suave ou forte (Dinâmicas), tocar frequências baixas ou altas (percepção de altura sonora), seleccionar o gerador de som (Timbre), estruturar as diferentes secções da obra, etc.⁷³

Ambos as obras foram compostas e executadas em tempo real através do MuDI. Na geração sonora não são utilizados *samples* nem qualquer sistema de som indeferido. Qualquer movimento realizado pelo intérprete resulta numa reacção sonora imediata.

No concerto, o computador portátil fica escondido da audiência e o MuDI é tocado no meio da sala. Desta maneira, desvia-se a atenção sobre o intérprete e consequentemente confere-se a devida proeminência ao som e às imagem projectadas. Por outro lado, o intérprete ao encontrar-se no centro da sala, fica numa posição privilegiada, em termos de auscultação sonora e de observação frontal das imagens.

A Figura 103 e a Figura 104 representam as interfaces de utilizador projectadas durante as interpretações.

⁷³ Estes aspectos composicionais podem ser visionados em <http://www.youtube.com/watch?v=UzYW1cHMFDM> ou através do DVD em anexo, no directório MuDI, com o nome *sonification – films for music\$2_aplicacao.mov*

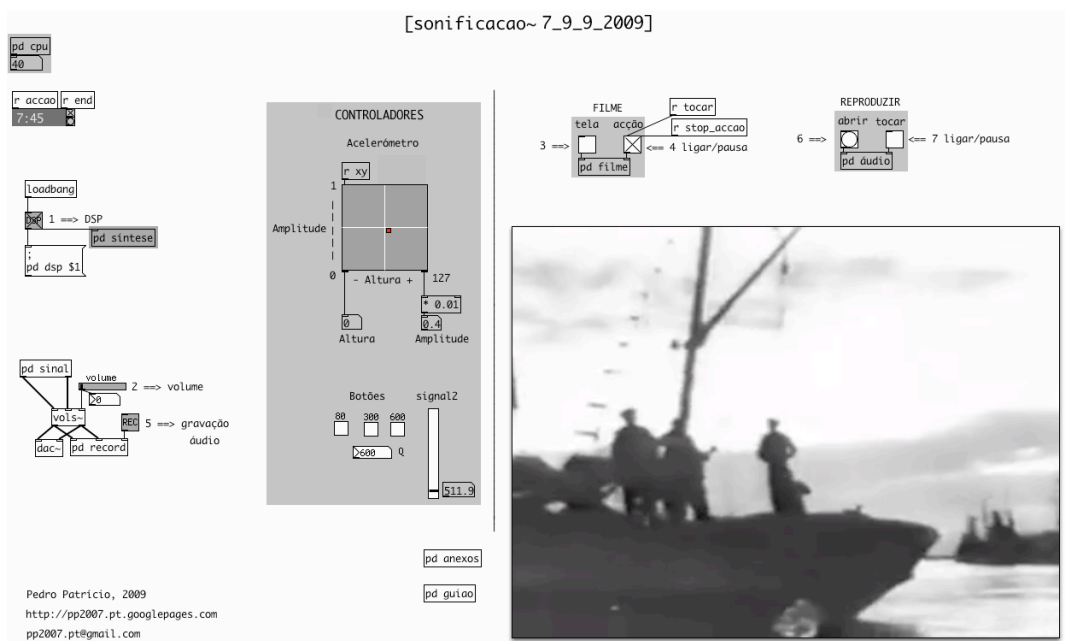


Figura 103. Interface de utilizador *sonificação - filmes para música\$1* (MuDI versão 1.0).⁷⁴

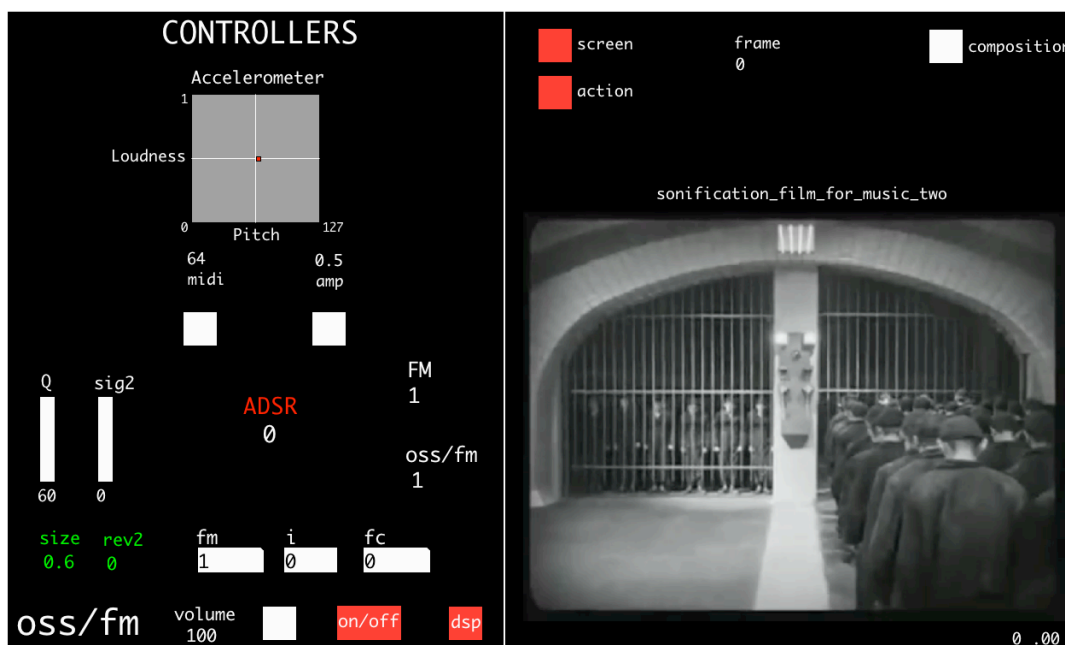


Figura 104. Interface de utilizador *sonification - films for music\$2* (MuDI versão 1.1).⁷⁵

⁷⁴ A operacionalização da interface de utilizador pode ser visualizada no DVD em anexo, no directório MuDI, com o nome *sonificação - filmes para música\$1_aplicacao.mov*

⁷⁵ A operacionalização da interface de utilizador pode ser visualizada no DVD em anexo, no directório MuDI, com o nome *sonification - films for music\$2_aplicacao.mov*

CONCLUSÕES E DIRECÇÕES FUTURAS

Nesta dissertação propusemos a construção de uma fonte sonora baseada no Som de Shepard, com o propósito de a utilizarmos como gerador de som, quer na concepção de um instrumento digital multimédia, quer em composição musical. A fonte sonora proposta foi designada de Chromasom, pelo facto, de ser um som que preserva a altura tonal, ou seja, a posição do som dentro da oitava.

O desafio de aplicar uma envolvente espectral alternativa ao Som de Shepard, tanto quanto sabemos nunca antes foi usada para reproduzir a Ilusão Sonora da Escala de Shepard foi alcançado.

Ao contrário das envolventes espectrais de configuração simétrica utilizadas por outros autores (Subcapítulo IX.2), em que as componentes das extremidades dos sons possuem sempre as amplitudes mais baixas e as componentes centrais as amplitudes mais altas, na envolvente espectral alternativa as amplitudes mais altas encontram-se sempre nas componentes mais à direita do espectro sonoro. Desta maneira, demonstrámos que é possível, não só reproduzir o Som de Shepard, como a Ilusão Sonora da Escala de Shepard através de uma envolvente espectral não simétrica.

A envolvente espectral alternativa concebida (baseada na inversão da curva de igual percepção de intensidade sonora de 40 *phons* de Fletcher e Munson) foi designada de Envolvente Espectral Alternativa Auto-Generativa (EEAAG). É auto-generativa, no sentido, em que permite preservar a igual percepção de intensidade sonora de qualquer frequência introduzida na sua variável. Ou seja, qualquer frequência, cuja intensidade sonora caia ao longo da curva de 40 *phons*, será ouvida com a mesma intensidade sonora de um som puro de 1000 Hz (frequência de referência). Assim, em termos psicoacústicos, em relação às frequências, vemos automaticamente compensadas as nossas diferenças de

sensibilidade auditiva. Consequentemente, no processo de composição musical, no que respeita aos sons sinusoidais, obtemos um maior equilíbrio sonoro entre os registos graves, médios e agudos.

O facto do Chromasom ser um som ambíguo em termos de percepção de altura sonora, como acontece com o Som de Shepard (Deutsch 2008, p. 589; Shepard 1999, p. 158) oferece outras possibilidades de geração som e de escuta musical. Provavelmente, conseguiremos definir a nota musical do som, mas teremos muita dificuldade em determinar a oitava a que pertence.

O Chromasom reproduziu a Ilusão Sonora da Escala de Shepard, mesmo quando aplicado à melodia que criámos (Figura 77). Por essa razão, decidimos dar à melodia o nome de Ilusão Sonora da Melodia Perpétua, porque quando é escutada cria a sensação de perpetuidade sonora, como acontece na Ilusão Sonora da Escala de Shepard. Desta maneira, demonstramos também, que é possível reproduzir a Ilusão Sonora da Escala de Shepard através de uma melodia constituída por uma sequência de saltos intervalares não regulares. Consequentemente, estamos a contribuir para expandir as possibilidades musicais da utilização da respectiva ilusão em composição musical.

Na nossa opinião, nos trabalhos musicais compostos durante a elaboração desta dissertação sobretudo, na obra musical intitulada “*Perpetual Melody - contrasting moments*”, o Chromasom demonstrou ser uma fonte sonora timbricamente interessante e musicalmente flexível. Flexível, no sentido, em que funciona sonora e musicalmente, quer com tempos lentos ou rápidos, quer com diferentes tipos de envolventes de amplitude e em todos os registos sonoros, desde os sons mais graves até aos sons mais agudos. Por essa razão, o Chromasom pode ser um gerador de som interessante e bem sucedido para compor música digital.

Em 1964, Shepard criou o Som de Shepard e a Ilusão Sonora da Escala de Shepard. Em 1969, Risset reproduziu a Ilusão Sonora da Escala de Shepard através de um glissando (Glissando Perpétuo). Nós, ao reproduzimos a Ilusão Sonora da Escala de Shepard através de uma melodia constituída por uma sequência de saltos intervalares não regulares (Ilusão Sonora da Melodia Perpétua), procurámos também dar o nosso contributo em vários domínios, quer ao nível (1) da acústica, através da aplicação ao Som de Shepard de uma envolvente espectral alternativa não simétrica, tanto quanto sabemos nunca antes foi usada para reproduzir a Ilusão Sonora da Escala de Shepard, (2) do esclarecimento da construção e mecanismo de funcionamento da ilusão, ao estudá-la, implementá-la em *Pure Data*, submetê-la a várias experiências auditivo-musicais, para

além da construção de aplicações com interfaces de utilizador, que ficarão disponíveis para a comunidade, quer através do DVD que acompanha a nossa dissertação, quer através da Internet, via descarregamento livre de conteúdos, (3) da aplicação das ilusões sonoras, ao utilizarmos esta ilusão numa melodia constituída por uma sequência de saltos intervalares não regulares (4) da composição musical, por termos ampliado a utilização musical da Ilusão Sonora da Escala de Shepard, por um lado, através da aplicação da respectiva ilusão sonora a uma melodia, do qual resultou a Ilusão Sonora da Melodia Perpétua e, por outro lado, na utilização da respectiva melodia como conteúdo musical exclusivo no trabalho intitulado “*Perpetual Melody – contrasting moments*”.

No que concerne ao trabalho desenvolvido no âmbito da interpretação musical com dispositivos electrónicos móveis gostaríamos de acrescentar o seguinte: o instrumento digital multimédia apresentado (MuDI) pertence a uma realidade musical, onde dispositivos electrónicos móveis comuns, como os PDAs e os telemóveis são reposicionados das suas funções originais e adaptados a ambientes artísticos com a finalidade de se transformarem em controladores ou instrumentos musicais.

O Mudi foi concebido para compor e tocar música para filmes em tempo real. Esta é a principal razão pela qual o MuDI permite o visionamento das imagens do filme, a gravação e a reprodução da composição musical em sincronização com o filme seleccionado.

Através do MuDI é possível obter automaticamente, para além de uma partitura vídeo, a composição musical e um ficheiro audiovisual da interpretação. Para obter o ficheiro audiovisual, basta importar o ficheiro áudio gravado durante a interpretação para o ficheiro QuickTime que contém o filme projectado. Através deste procedimento é possível criar, em tempo real, a sonoplastia ou compor a música de qualquer filme.

A projecção da interface de utilizador permite ao intérprete e à audiência acompanhar, não só as imagens do filme, mas também a relação entre os gestos realizados e o som gerado.

A abordagem adoptada na construção do Módulo Som (Figura 84) permite que os geradores de som sejam reprogramados e adaptados a qualquer trabalho musical. Por exemplo, se pretendermos obter o som de um *Theremin*, basta programar um *subpatch* de síntese de som por FM em *Pure Data* e conectá-lo ao *subpatch* “*pd synthesis*”. A partir desse momento, o novo gerador de som fica pronto a ser usado.

O modo de operação do MuDI foi pensado de maneira a manter a sua simplicidade de manuseamento. Paralelamente, o método de calibração implementado permitiu ultrapassar a falta de um ponto de referência fiável do acelerómetro. Acreditamos que a simplicidade, a fiabilidade e a estabilidade do sistema, aliada a um controlo efectivo sobre o som, permitirá melhorar a confiança do compositor/intérprete e consequentemente alcançar melhor expressão musical.

Aplicar o Chromasom a outras técnicas de síntese de som digital (e.g., síntese por modulação em anel, recirculação de *delay*, etc.) com o propósito de encontrar novos timbres para compor novas musicas será uma das direcções futuras a perseguir.

Em relação ao MuDI, a subsequente fase deste trabalho passará por continuar a testar as suas principais características, no sentido, de o tornar ainda mais robusto, expressivo e de fácil manuseamento. Para além do ruído branco e do Chromasom pretendemos introduzir novos geradores de som. Estes, serão organizados de maneira a serem automaticamente seleccionados pelo utilizador, através da introdução de mais controladores na aplicação do dispositivo electrónico móvel. Desta maneira, o compositor/intérprete terá à sua disposição um maior leque de alternativas timbricas para compor e tocar novas musicas em tempo real.

Em termos de composição musical pensamos criar um novo ciclo denominado por *Shepard Music*, com o objectivo de testar e expandir ainda mais a possibilidade de escuta musical apresentada nesta dissertação. Para esse fim, utilizaremos o Chromasom como gerador de som aplicado a um sintetizador polifónico digital programado em *Pure Data*. O sintetizador será controlado através de um teclado MIDI. Para alcançar este propósito começaremos por seleccionar, tocar e gravar algumas obras musicais, de vários compositores com o objectivo de observar o efeito produzido resultante da introdução de um novo instrumento no reportório já constituído.

Pretendemos também, desenvolver novas técnicas de composição, a partir dos exemplos sonoros existentes no CD, que acompanha o trabalho de Bregman (1996) intitulado “*Demonstrations of Auditory Scene Analysis: The Perceptual Organization of Sound. Audio Compact Disk*”. Na abordagem a esta temática começaremos por realizar uma selecção criteriosa dos exemplos sonoros de Bregman, que ofereçam inicialmente um maior potencial musical por explorar. De seguida, os exemplos seleccionados seriam reproduzidos através de processos de síntese de som digital e aplicados, como novas

técnicas à composição musical.

As direcções futuras apresentadas respondem à necessidade e ao desejo prementes de querermos aprender e produzir mais conhecimento nos domínios do som e da música, pois como afirma Maria Gabriela Llansol (1991, p. 99),

“o som cria o ouvido, o ouvido faz o cérebro, o cérebro concebe a existência do homem...”⁷⁶

⁷⁶ Llanson, G. M. (1991). Um beijo dado mais tarde. Ed. Rolim, 2.^a edição.

BIBLIOGRAFIA

Barkenecht, F. (2007). Beginner's Guide to the FFT-objects in Pd. Extraído 20 Janeiro 2011 em <http://footils.org/2007/02/20/beginners-guide-fft-objects-pd/>

Benson, D. (2008). Music: A Mathematical Offering. Web version. Extraído 2 Novembro 2010 em <http://www.maths.abdn.ac.uk/~bensondj/html/music.pdf>

Bregman, A. S. (1990). Auditory Scene Analysis: The Perceptual Organization of Sound. Cambridge, Massachusetts: The MIT Press.

Bregman, A. S. (1993). Thinking in sound: the cognitive psychology of human audition. Ed. S. McAdams e E. Bigand. Oxford University Press, pp. 10-36.

Bregman, A. S., Ahad, P. (1996). Demonstrations of Auditory Scene Analysis: The Perceptual Organization of Sound. Audio Compact disk. Mit Press.

Burns, M. E. (1981). Circularity in relative pitch judgments for inharmonic complex tones: The Shepard demonstration revisited, again. *Perception e Psychophysics*, 30 (5), pp. 467-472.

Cheng, C. I., Wakefield, G. H. (2001). Introduction to Head Related Transfer Functions (HRTFs): Representations of HRTFs in Time, Frequency and Space. *Journal of the Audio Engineering Society*, New York, 49 (4), pp. 231-249.

Cheveigné, A. (2004). Pitch Perception Models in Pitch. Springer-Verlag. Ed. Plack e A. Oxenham.

Cheveigné, A. (2010). Pitch Perception. Ed. Plack e A. Oxenham. Oxford University Press, pp. 71-104.

Chowning, J. (1973). The Synthesis of Complex Audio Spectra by Means of Frequency Modulation. *Journal of the Audio Engineering Society* 21 (7), pp. 526-534.

Chowning, J. (1999). Perceptual Fusion and Auditory Perspective. Music, Cognition, and Computerized Sound: An Introduction to Psychoacoustics. The MIT Press, Cambridge, Massachusetts, London, England, Ed. Perry R. Cook, pp. 261-276.

Cook, R. P. (1999). *Music, Cognition, and Computerized Sound: An Introduction to Psychoacoustics*. Cambridge, MA: MIT Press.

DeCasper, A., Fifer, W. (1980). Of Human Bonding: Newborns Prefer their Mothers' Voices. *Science*, New Series, 208 (4448). Published by: American Association for the Advancement of Science Stable, pp. 1174-1176.

Dekel, G., Dekel, A. (2008). MoGMI: Mobile Gesture Music Instrument. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*, Vienna, Austria, pp. 6-10. Extraído 4 Outubro 2010 em <http://www.mobilemusicworkshop.org/>

Deutsch, D. (1972). Octave Generalization and Tune Recognition. *Perception and Psychophysics*, 11, pp. 411-412.

Deutsch, D. (1974). Octave Illusion. *Nature*, 251, pp. 307-309.

Deutsch, D. (1981). The Octave Illusion and Auditory Perceptual Integration. Ed. J. V. Tobias e E. D. Schubert, *Hearing Research and Theory*, 1, New York: Academic Press, pp. 99-142.

Deutsch, D., Moore F. R., Dolson, M. (1986). The Perceived Height of Octave-Related Complexes. *Journal of Acoustic Society of América*, 80 (50), pp. 1346-1353.

Deutsch, D. (1987). The Tritone Paradox: Effects of Spectral Variables. *Perception & Psychophysics*, 41, pp. 563-575.

Deutsch, D. (1992). Some New Pitch Paradoxes and Their Implications. The Royal Society, London, England, 366, pp. 391-397.

Deutsch, D. (1999). Grouping Mechanisms in Music. *The Psychology of Music*. Second Edition, San Diego: Academic Press, pp. 299-348.

Deutsch, D., Dooley, K., Henthorn, T. (2008). Pitch Circularity from Tones Comprising Full Harmonic Series. *Journal of Acoustic Society of America*, 124 (1), pp. 589-597.

Deutsch, D. (2010). The Paradox of Pitch Circularity. *Acoustic Today*, pp. 8-15.

Duda, R. O. (1993). Modeling Head Related Transfer Functions. In *Proceedings of the 27th Conference on Signals, Systems & Computers*, Asilomar, 2, pp. 996-1000.

Emanuel, D., Maroonroge, S., Letowski, T. (2009). Auditory Function. Ed. C. E. Rash, M. B. Russo, T. R. Letowski, e E. T. Schmeisser, *Helmet-mounted displays: Sensation, perception and cognition issues*. Ft. Rucker, AL: U. S. Army Aeromedical Research Laboratory, pp. 307-334.

Essl, G., Rohs, M. (2007). Shamus – a Sensor-based Integrated Mobile Phone Instrument. In Proceedings of the International Computer Music Conference (ICMC), Copenhagen. Extraído 4 Outubro 2010 em <http://www.eecs.umich.edu/~gessl/>

Essl, G., Rohs, M. (2009). Interactivity for Mobile Music-Making. *Organised Sound*, 14 (2), pp. 197-207.

Essl, G., Rohs, M., Kratz, S. (2010). Use the Force (or something) – Pressure and Pressure – Like Input for Mobile Music Performance. In Proceedings of the International Conference on New Interfaces for Musical Expression (NIME), Sydney, Australia, pp. 182-185. Extraído 1 Novembro 2010 em <http://www.educ.dab.uts.edu.au/nime/PROCEEDINGS/>

Essl, G., Wang, G., Rohs, M. (2008). Developments and Challenges Turning Mobile Phones into Generic Music Performance Platforms. In Proceedings of the International Conference on New Interfaces for Musical Expression (NIME), Vienna, Austria, pp. 11-14. Extraído 4 Outubro 2010 em <http://recherche.ircam.fr/equipes/temps-reel/nime06/>

Farnell, A. (2008). Designing Sound. Applied Scientific Press, London, England. ISBN-13: 978-0-9560886-0-4

Féron, F. X. (2006). Des illusions auditives aux singularités du son et de la perception: l’impact de la psychoacoustique et des nouvelles technologies sur la création musicale au XXe siècle. Tese de Doutorado, Paris IV, Sorbonne, Paris.

Fugiel, B. (2011). Waveform Circularity From Added Sawtooth And Square Wave Acoustical Signals. *Music Perception*, 28 (4), pp. 415-423.

Gaye, L., Holmquist, E. L., Behrendt, F., Tanaka, A. (2006). Mobile Music Technology: Report on an Emerging Community. In Proceedings of the International Conference on New Interfaces for Musical Expression (NIME), Paris, France, pp. 22-25. Extraído 1 Novembro 2010 em <http://recherche.ircam.fr/equipes/temps-reel/nime06/>

Geiger, G. (2003). Pda: Real Time Signal Processing and Processing and Sound Generation on Handheld Devices. In Proceedings of the International Computer Music Conference (ICMC), Singapore, pp. 1-4. Extraído 4 Outubro 2010 em http://quod.lib.umich.edu/cgi/t/text/text-idx?c=icmc;idno=bbp2372.2003.*

Geiger, G. (2006). Using the Touch Screen as a Controller for Portable Computer Music Instruments. In Proceedings of the International Conference on New Interfaces for Musical Expression (NIME), Paris, France, pp. 61-64. Extraído 1 Novembro 2010 em http://recherche.ircam.fr/equipes/temps-reel/nime06/proc/nime2006_061.pdf

- Goldstein, E. B. (1981). The Ecology of J. J. Gibson's Perception. *Leonardo*, 14 (3), pp. 191-195.
- Goolkasian, P., Woodberry, C. (2010). Priming Effects with Ambiguous Figures. *Attention, Perception and Psychophysics*, 72 (1), pp. 168-178.
- Gregory, L. R. (1968). Perceptual Illusions and Brain Models. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 171 (1024). A Discussion on the Logical Analysis of Cerebral Functions (Dec. 13, 1968), pp. 279-296.
- Gregory, L. R. (1997). Visual Illusions Classified. *Trend in Cognitive Sciences*. Vol. 1 (5), pp. 190-194.
- Gregory, L. R. (1997). Knowledge in Perception and Illusion, 352, Department of Psychology, University of Bristol, UK, Phil. Trans. R. Soc. Lond. B, pp. 1121 – 1128
- Gregory, L. R. (1998). Eye and Brain: The Psychology of Seeing. Fifth Edition, Princeton University Press.
- Hall, E. D. (1990). Musical Acoustics. California State University, Sacramento, Brooks/Cole Publishing Company, Pacific Grove, California, 2.^a Ed.
- Helmholtz, H. (1895). On the Sensations of Tone as a Physiological Basis for the Theory of Music. Longmans, Green, and Co, 3.^a Ed.
- Hofstadter, D. R. (1979). Godel, Escher, Bach: an Eternal Golden Braid. New York, Basic Books Inc.
- Kanizsa, G. M. (1955). Quasi-percettivi in Campi con Stimolazione Omogenea. *Rivista di Psicologia*, 49, (I). Istituto di Psicologia dell'Università di Trieste, pp. 7-30.
- Kester, W. (2000). Mixed-Signal and DSP Design Techniques. Section 5, Analog Devices, Inc. USA. ISBN-0-916550-23-0
- Kreidler, J. (2007). Loadbang. Programming Electronic Music in Pure Data. Extraído 20 Outubro 2010 em <http://www.pd-tutorial.com/english/index.html>
- Levitin, D. J. (2006). This is Your Brain on Music: The Science of a Human Obsession. New York, Dutton, Penguin.
- Levitin, D. J. (2007). Psychology of Music. Darity's International Encyclopedia of Social Sciences, 2nd Edition, Ed. MacMillan.
- Levitin, D. J. (2008). The Illusion of Music. New Scientist Magazine, Music Special, 2644, pp. 34-37.
- Licklider, J. C. R. (1951). A Duplex Theory of Pitch Perception. *Experimenta* 7 (4), pp. 128-134.
- Loy, G. (2006). Musimathics 1: The Mathematical Foundations of Music.

Cambridge, Massachusetts, The MIT Press.

Maroonroge, S., Emanuel, D., Letowski, T. (2009). Basic Anatomy of the Hearing System. Ed. Inc. E. Rash, M. B. Russo, T. R. Letowski, and E. T. Schmeisser. *Helmet-mounted displays: Sensation, perception and cognition issues*. Ft. Rucker, AL: U. S. Army Aeromedical Research Laboratory, pp. 279-306.

Mathews, M. V. (1963). The digital computer as a musical instrument. *Science*, 142, pp. 553-557.

Mathews, M. (1977). Analysis and Synthesis of Timbre. Royal Swedish Academy, 17, pp. 4-18. ISBN 91-85428-03-5.

Mathews, M. (1999). The Auditory Brain. Music, Cognition, and Computerized Sound: An Introduction to Psychoacoustics. The MIT Press, Cambridge, Massachusetts, London, England, Ed. Perry R. Cook, pp. 261-275.

Mathews, M. (1999). The Ear and How it Works. Music, Cognition, and Computerized Sound: An Introduction to Psychoacoustics. The MIT Press, Cambridge, Massachusetts, London, England, Ed. Perry R. Cook, pp. 1-10.

Mathews, M. (1999). What is Loudness? Cognition, and Computerized Sound: An Introduction to Psychoacoustics. The MIT Press, Cambridge, Massachusetts, London, England, Ed. Perry R. Cook, pp. 261-275.

Mcgurk, H., MacDonald, J. (1976). Hearing Lips and Seeing Voices. *Nature*, 264 (5588), pp. 746-748.

Nakajima, Y., Tsumura, T., Matura, S., Minami, H., Teranishi, R. (1988). Dynamic Pitch Perception for Complex Tones Derived from Major Triads. *Music Perception*. University of California, 6 (1), pp. 1-20.

Necker, L. A. (1832). Observations on Some Remarkable Optical Phaenomena Seen in Switzerland; and On An Optical Phenomenon, Which Occurs on Viewing a Figure of a Crystal or Geometrical Solid. The London and Edinburgh Philosophical Magazine and *Journal of Science*, 1 (5), pp. 329-337.

Nudds, M. (2007). Auditory Perception and Sounds: New Philosophical Essays. Oxford University Press.

Oh, J., Herrera, J., Bryan, J. N., Dahl, L., Wang, G. (2010). Evolving the Mobile Phone Orchestra. In Proceedings of the International Conference on New Interfaces for Musical Expression (NIME), Sydney, Austrália, pp. 82-87. Extraído 1 Novembro 2010 em <http://www.educ.dab.uts.edu.au/nime/PROCEEDINGS/>

Palmer, S. E., Brooks, J. L., Nelson, R. (2003). When Does Grouping Happen?. *Acta Psychologica*, 114, pp. 311-330.

Palmer, S. E. (2002). Perceptual Organization in Vision. *Steven's Handbook of Experimental Psychology*, Third Edition, 1, pp. 177-234.

Pierce, J., Schubert, E. D. (1987). High Level Factors and the Musical Saliency of Auditory Phenomena. Center for Computer Research in Music and Acoustic (CCRMA). Department of Music. Standford University, Standford, Califórnia. Report n.º STAN-M-41, pp. 1-14.

Pierce, J. (1999). Consonance and Scales. *Music, Cognition, and Computerized Sound: An Introduction to Psychoacoustics*. The MIT Press, Cambridge, Massachusetts, London, England, Ed. Perry R. Cook, pp. 167-185.

Pierce, J. (1999). Introduction to Pitch Perception. *Music, Cognition, and Computerized Sound: An Introduction to Psychoacoustics*. The MIT Press, Cambridge, Massachusetts, London, England, Ed. Perry R. Cook, pp. 57-70.

Pierce, J. (1999). Sound Waves and Sine Waves. *Music, Cognition, and Computerized Sound: An Introduction to Psychoacoustics*. The MIT Press, Cambridge, Massachusetts, London, England, Ed. Perry R. Cook, pp. 37-56.

Plack, C. J. (2004). Auditory perception. Psychology Press Ltd, pp. 1-19.

Pollack, I. (1978). Decoupling of Auditory Pitch and Stimulus Frequency: The Shepard Demonstration Revisited. *Journal of the Acoustical Society of America*, 63, pp. 202-206.

Puckette, M. (1996). Pure Data: Another Integrated Computer Music Environment. In *Proceedings of the Second Intercollege Computer Music Concerts*, Tachikawa, Japan. Extraído 17 Junho 2009 em <http://www-crcra.ucsd.edu/~msp/publications.html>

Puckette, M. (2002). Max at Seventeen. *Computer Music Journal*, 26(4), pp. 31-43.

Puckette, M. (2007). *The Theory and Technique of Electronic Music*. University of California, San Diego, USA. ISBN: 978-981-270-077-3.

Risset, J. C. (1971). Paradoxes de Hauteur: Le Concept de Hauteur Sonore n'est pas le Meme Pour Tout le Monde, *Seventh International Congress on Acoustic*, 20 S 10, Budapest, pp. 613-616.

Risset, J. C. (1977). Paradoxes de Hauteur. Communication au Symposium IRCAM sur la psychoacoustique musicale. University Lyon II, France, pp. 1-6.

Risset, J. C. (1992). Le Son Numérique: Une Acoustique Affranchie de la Mécanique? *Journal de Physique IV*.

Risset, J. C. (1994). Sculpting Sounds with Computers: Music, Science, Technology. *Leonardo*, 27 (3). Art and Science Similarities, Differences and Interactions. The Mit Press, pp. 257-261. Extraído 4 Janeiro 2011 em www.jstor.org/stable/1576064

Risset, J. C. (1997). Computer Music Why? Extraído 2 Outubro 2010 em http://www.utexas.edu/cola/insts/france-ut/_files/pdf/resources/risset_2.pdf

Risset, J. C. (1998). Rhythmic Paradoxes and Illusions: a Musical Illustration. In Proceedings of International Computer Music Conference (ICMC), Thessaloniki, Greece, pp. 7-10.

Risset, J. C. (2007). Fifty Years of Digital Sound for Music. In Proceedings of Sound and Music Computing Conference (SMC), Lefkada, Greece, pp. 3-7. Extraído 4 Outubro 2010 em <http://smc07.uoa.gr/SMC07%20Proceedings/SMC07%20Paper%201.pdf>

Risset, J. C. (2009). Max Mathews's Influence on (My) Music. *Computer Music Journal*, 33 (3), pp. 26-34.

Rohs, M., Essl, G., Roth, M. (2006). CaMus: Live Music Performance using Camera Phones and Visual Grid Tracking. Proceedings of the 6th International Conference on New Instruments for Music Expression (NIME), pp. 31-36. Extraído 31 January 2011 em http://recherche.ircam.fr/equipes/temps-reel/nime06/proc/nime2006_031.pdf

Scharine, A. A., Letowski, T. R., Cave, K. (2009). Auditory Perception and Cognitive Performance. Ed. C. E. Rash, M. B. Russo, T. R. Letowski, and E. T. Schmeisser, *Helmet-mounted displays: Sensation, perception and cognition issues*. Ft. Rucker, AL: U. S. Army Aeromedical Research Laboratory, pp. 391-490.

Scharine, A. A., Letowski, T. R. (2009). Auditory Conflicts and Illusions. Ed. C. E. Rash, M. B. Russo, T. R. Letowski, and E. T. Schmeisser. *Helmet-mounted displays: Sensation, perception and cognition issues*. Ft. Rucker, AL: U. S. Army Aeromedical Research Laboratory, pp. 579-598.

Schiemer, G., Havryliv, M. (2005). Pocket Gamelan: a Pure Data Interface for Mobile Phones. In Proceedings of the International Conference on New Interfaces for

Musical Expression (NIME), Vancouver, BC, Canadá, pp. 37-42. Extraído 4 Outubro 2010 em <http://www.nime.org/2005/proceedings.html>

Schiemer, G., Havryliv, M. (2006). Pocket Gamelan: Tuneable Trajectories for Flying Sources in Mandala 3 and Mandala 4. In Proceedings of the International Conference on New Interfaces for Musical Expression (NIME), Paris, France, pp. 37-42. Extraído 4 Outubro 2010 em <http://recherche.ircam.fr/equipes/temps-reel/nime06/>

Schroeder, R. M. (1986). Auditory Paradox Based on Fractal Waveform. *Journal of the Acoustical Society of America*, 79, pp. 186-188.

Shams, L., Kamitani, Y., Shimojo, S. (2002). Visual Illusion Induced by Sound. *Cognitive Brain Research*, 14, pp. 147-152.

Shepard, R. (1990). Mind Sights: Original Visual Illusions, Ambiguities, and Other Anomalies. With a Commentary on the Play of Mind in Perception and Art. New York: WH Freeman and Co.

Shepard, R. (1999). Pitch Perception and Measurement. Cognition, and Computerized Sound: An Introduction to Psychoacoustics. The MIT Press, Cambridge, Massachusetts, London, England, Ed. Perry R. Cook, pp. 149-165.

Shepard, R. (1999). Stream Segregation and Ambiguity in Audition. Cognition, and Computerized Sound: An Introduction to Psychoacoustics. The MIT Press, Cambridge, Massachusetts, London, England, Ed. Perry R. Cook, pp. 117-127.

Shepard, R. (1964) Circularity in Judgments of Relative Pitch. *Journal of the Acoustic Society of America*, 36 (12), pp. 2346-2353.

Slater, A. (2002). Visual Perception in the Newborn Infant: Issues and Debates. *Intellectica*, 34, pp. 57-76.

Smith, W. S. (1999). The Scientist and Engineer's Guide to digital Signal Processing. Second Edition, California Technical Publishing, P. O. Box, San Diego, CA 92150. Capítulos 8 e 12. Extraído 10 Dezembro 2010 em <http://www.dspguide.com>

Suzuki, Y., Takeshima, H. (2004). Equal-loudness Contours for Pure Tones. *Journal of Acoustic Society of America*, 116 (2), pp. 918-933.

Suzuki, Y. et al. (2003). Precise and Full-range Determination of Two-dimensional Equal Loudness Contours. Technical report, Tohoku University, Japan. Extraído 15 Novembro 2010 em <http://www.mp3-tech.org/programmer/docs/IS-01Y-E.pdf>

Tanaka, A. (2004). Malleable Mobile Music. In Adjunct Proceedings of the 6th International Conference on Ubiquitous Computing (Ubicomp), Nottingham, England.

<http://www.ubicomp.org/ubicomp2004/adjunct/demos/tanaka.pdf>

Ueda, K., Ohgushi, K. (1987). Perceptual Components of Pitch: Spatial Representation Using a Multidimensional Scaling Technique. *Journal of the Acoustic Society of America*, 82, pp. 1193-1200.

Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., Heydt, R. (2012). A Century of Gestalt Psychology in Visual Perception: I. Perceptual Grouping and Figure–Ground Organization. *Psychological Bulletin*. 138 (6) pp. 1172-1217.

Wang, G., Cook, R. P. (2004). On-the-fly Programming: Using Code As An Expressive Musical Instrument. In Proceedings of the International Conference on New Interfaces for Musical expression (NIME). Hamamatsu, Japan, pp. 138-143. Extraído 10 Junho 2009 em <http://www.nime.org/2004/>

Wang, G., Essl, G., Penttinen, H. (2008). MoPhO: Do Mobile Phones Dreams of Electric Orchestras? In Proceedings of the International Computer Music Conference (ICMC), Belfast, UK, pp. 24-29. Extraído 10 Junho 2009 em <https://ccrma.stanford.edu/~ge/publish/>

Warren, R. M. (2008). Auditory Perception: An Analysis and Synthesis. Cambridge University Press, New York. ISBN-13: 978-0-521-86870-9.

Warren, R. M., Warren, R. P. (1970). Auditory Illusions and Confusions. *Scientific American*, pp 30-36.

Warren, R. M., Gregory, R. L. (1958). An Auditory Analogue of the Visual Reversible Figure. *American Journal of Psychology*, 71, pp. 612–613.

Wong, G. (2007). Sound Level Meters. The Handbook of Noise and Vibration Control. Ed. Malcolm J. Crocker, New York, Chapter 38, pp. 455-464.

Wright, M. (2005). Open Sound Control: An Enabling Technology for Musical Networking. *Organised Sound*, 10 (3), pp. 193-201.

Sites

<http://www4.uwm.edu/Dept/APL/demonstrations.html>

Richard M. Warren homepage/Auditory Demonstrations (última consulta em 2010.04.11).

<http://www.simplypsychology.org/perception-theories.html>

(última consulta em 2009.05.01).

<http://www.cochlea.org/>

(última consulta em 2010.05.12).

http://www.daviddarling.info/encyclopedia/O/organ_of_Corti.html

The Internet Encyclopedia of Science (última consulta em 2010.05.12).

<http://plato.stanford.edu/entries/sounds/>

Standford Encyclopedia of Philosophy, Sounds, First Published Thursday July 14, 2005 (última consulta em 2010.06.30).

<http://www.priberam.pt/>

Dicionário de Língua Portuguesa On-line (última consulta em 2010.06.30).

<http://www.infopedia.pt/>

Dicionário de Língua Portuguesa On-line (última consulta em 2010.06.30).

<http://www.oed.com/services/>

On-line Oxford English Dictionary (última consulta em 2010.06.30).

http://www.kayelaby.npl.co.uk/general_physics/2_4/2_4_1.html

The speed and attenuation of sound (última consulta em 2010.06.02).

GLOSSÁRIO

A

Altura espectral – *Tone height*

Altura tonal – *Tone chroma*

análise de Fourier – *Fourier analysis*

C

Curvas de igual percepção de intensidade sonora – *Equal loudness curves*

Constância de percepção de intensidade sonora – *Loudness constancy*

Constância de tamanho – *Size constancy*

D

DFT inversa – *Inverse DFT*

I

Ilusão Sonora da Escala de Shepard – *Shepard Scale Illusion*

Ilusão Visual das pernas do elefante de Shepard – *Shepard's elephant legs visual illusion*

Inversão da transformada discreta de Fourier – *Inverse discrete Fourier transform* (IDFT)

P

Percepção de altura sonora – *Pitch*

Percepção de intensidade sonora – *Loudness*

Processamento de som digital – *Digital Sound Processing* (DSP)

S

Som de Shepard – *Shepard Tone*

T

transformada de Fourier – *Fourier transform*

transformada discreta de Fourier – *discrete Fourier transform* (DFT)

transformada de Fourier em tempo discreto – *discrete time Fourier transform*

transformada rápida de Fourier – *fast Fourier transform* (FFT)

transformada rápida de Fourier real – *real fast Fourier transform* (RFFT)

ANEXOS

ANEXO 1: PUBLICAÇÕES, CONFERÊNCIAS

Lista de Artigos do Autor Publicados durante o Doutorado em Conferências com Escrutínio Científico

Patrício, P. (2012). From the Shepard Tone to the Perpetual Melody Auditory Illusion. Actas da 9.^a Conferência Internacional *Sound and Music Computing* (SMC), Universidade de Aalborg, Copenhaga, Dinamarca, pp. 5-10. <http://smc2012.smcnetwork.org/>

Patrício, P. (2012). MuDI – Multimédia Digital Instrument for Composing and Performing Digital Music for Films in Real-time. Actas da 12.^a Conferência Internacional *New Interfaces for Musical Expression* (NIME), Universidade de Michigan, Ann Arbor, Michigan, EUA, pp. 264-267. <http://www.eecs.umich.edu/nime2012/index.htm>

Patrício, P. (2011). From the Shepard Tone to the Perpetual Simple Melody Auditory Illusion. Actas do 3.º INForum - Simpósio de Informática, Coimbra, Portugal, pp. 521-532. <http://inforum.org.pt/INForum2011>

Patrício, P. (2009). A Música no Cinema Mudo e o Instrumento Musical Digital. Actas do 12.º Simpósio Brasileiro de Computação Musical (SBCM), Recife, Pernambuco, Brasil, pp. 151 - 162. <http://compmus.ime.usp.br/sbcm/2009/english/index.html>

Lista de Conferências com Escrutínio Científico Realizadas pelo Autor durante o Doutoramento

Patrício, P. From the Shepard Tone to the Perpetual Melody Auditory Illusion. 9.^a Conferência Internacional Sound and Music Computing (SMC), Universidade de Aalborg, Copenhaga, Dinamarca, 11-14 Julho 2012. <http://smc2012.smcnetwork.org/>

Patrício, P. MuDI – Multimédia Digital Instrument for Composing and Performing Digital Music for Films in Real-time. Actas da 12.^a Conferência Internacional *New Interfaces for Musical Expression* (NIME), Universidade de Michigan, Ann Arbor, Michigan, EUA, 21-23 Maio 2012. <http://www.eecs.umich.edu/nime2012/index.htm>

Patrício, P. MuDI - Multimedia Digital Instrument for Composing and Performing Digital Music for Films in Real-time. 2.º Fórum Itinerários Musicais - Música e Gesto, Lisboa, Portugal, 28-30 Outubro 2011. http://cesem.fcsh.unl.pt/events/2ndforum-musical-itineraries-music-and-gesture?set_language=en

Patrício, P. From the Shepard Tone to the Perpetual Simple Melody Auditory Illusion. 3.º INForum - Simpósio de Informática, Coimbra, Portugal, 8-9 Setembro 2011. <http://inforum.org.pt/INForum2011>

Patrício, P. Ilusões Sonoras: erros dos sentidos, verdades da percepção. Jornadas de Doutoramento, Escola das Artes, UCP, Porto, Portugal, 2 Outubro 2010. http://artes.ucp.pt/jornadas_phd/programa.html

Patrício, P. A Música no Cinema Mudo e o Instrumento Musical Digital. 12.º Simpósio Brasileiro de Computação Musical (SBCM), Recife, Pernambuco, Brasil, 7-9 Setembro 2009. <http://compmus.ime.usp.br/sbcm/2009/english/index.html>

ANEXO 2: PRODUÇÃO ARTÍSTICA

Obras Musicais Compostas e Estreadas Durante o Doutorado

Patrício, P. *mudice\$2 – a irreversível*, suporte fixo, *Cadavre Exquis* electroacústico, Festival Música Viva, Orquestra de Altifalantes, Instituto Franco-Português, Lisboa, Novembro 2014.

Patrício, P. *mudice\$2 – a irreversível*, suporte fixo, *Center for Experimental Music and Intermedia* (CEMI), *Cadavre Exquis, Portuguese Composers of the 21st Century*, University of North Texas, USA, Outubro 2014.

Patrício, P. *mudice\$2 – a irreversível*, suporte fixo, *Cadavre Exquis* electroacústico, Festival Música Viva, Orquestra de Altifalantes & Ars Nova Ensemble, Instituto Franco-Português, Lisboa, Setembro 2010.

Patrício, P. *mudice\$3 – a perpétua*, suporte fixo, *Sound Walk*, Centro cultural de Belém, Passeio Pedonal, Festival Música Viva, Lisboa, Setembro 2010.

Patrício, P. *mudice\$2 – a irreversível*, suporte fixo, *Cadavre Exquis* electroacústico, Miso Music Portugal 25 anos, Instituto Franco-Português, Lisboa, Abril 2010.

Patrício, P. *sonificacao - filmes para musica\$1, Laptop e iPod touch*, 12.º Simpósio Brasileiro de Computação Musical, Recife, Pernambuco, Brasil.

Concertos Realizados Durante o Doutorado

2014.11.29, Festival de Música Viva 2014, *mudice\$2 – a irreversível*, *Cadavre Exquis* electroacústico, Orquestra de Altifalantes, O'culto da Ajuda, Lisboa.

2014.10.02, *mudice\$2 – a irreversível*, *Cadavre Exquis, Portuguese Composers of the 21st Century*, University of North Texas, USA.

2010.09.25, *mudice\$2 – a irreversível*, *Cadavre Exquis*, Festival de Música Viva

2010, Orquestra de Altifalantes & Ars Nova Ensemble, Instituto Franco-Português, Lisboa.

2010.09.16 a 19, *mudice\$3 – a perpétua, Sound Walk*, Centro Cultural de Belém, Passeio Pedonal, Festival Música Viva 2010, Lisboa.

2010.04.15, *mudice\$2 – a irreversível, Cadavre Exquis*, Miso Music Portugal 25 anos, Instituto Franco-Português, Lisboa.

2009.09.08, *sonificacao - filmes para musica\$1*, 12.º Simpósio Brasileiro de Computação Musical, Recife, Pernambuco, Brasil.

2009.05.15, *como quem respira*, Lançamento da BitRadio Records, PinUp Bar, Porto.

2009.04.22, Cerimónia de Abertura: (De) Criando... à sombra de Oliveira/Formação Variável de Laptops do Citar, 6º Festival Black and White, UCP – Escola das Artes, Porto.

2009.04.25, Cerimónia de Encerramento: Carlos Zíngaro + Formação Variável de Laptops do Citar, 6º Festival *Black and White*, UCP – Escola das Artes, Porto.

2009.01.30, Isto não é um Concerto, Orchestrutopica, Centro Cultural de Belém – Lisboa.

2008.12.12, O tal Encontro, Escola das Artes, Universidade Católica Portuguesa - Porto.

Publicação em CD

2014.11.27, *mudice\$2 – a irreversível, Cadavre Exquis, Portuguese Composers of the 21st Century*, Miso Records, mcd 036.13, Lisboa, Portugal, 1.ª Edição.

Divulgação Radiofónica

2015.01.03, *mudice\$2 – a irreversível, Cadavre Exquis, Portuguese Composers of the 21st Century*, Música Hoje, Antena 2.

ANEXO 3: LISTAGEM DO CONTEÚDO DO DVD

O conteúdo do DVD encontra-se distribuído por vários directórios. Por ordem alfabética, os directórios e os conteúdos agregados são:

Directório experiência auditivo-musical:

- Aplicacao.pd
- ISES_reproducao.wav
- ISMP.wav
- Shepard_original.wav

Directório MuDI:

- MuDI_itinerarios_musicais_2011.mov
- MuDI_nime_2012.mov
- MuDI_promoção.mp4
- MuDI_sbcm_2009.avi
- sonificação – música para filmes\$1_aplicacao.mov
- sonificação – música para filmes\$1.mov
- sonification – films for music\$2_aplicacao.mov
- sonification – films for music\$2.mov

Directório MuDICE:

- mudice\$1 – música de uma tecla só (a pretensiosa)_interface.mov
- mudice\$1 – música de uma tecla só (a pretensiosa).wav

- mudice\$2 – a irreversível.wav
- mudice\$3 – a perpétua_interface.mov
- mudice\$3 – a perpétua.wav
- perpetual melody – contrasting moments.mov
- perpetual melody – contrasting moments.wav
- perpetual_melody_contrasting_moments.pd

Directório patches:

- 00.leia-me.pd
- 01.efeito_sonoro_ambiguo.pd
- 02.tone_height.pd
- 03.equacao_2.pd
- 04.extracao_automatica_frequencias_individuais.pd
- 05.notas_musicais.pd
- 06.equacoes_3_e_4.pd
- 07.extracao_automatica_amplitudes_individuais.pd
- 08.sintese_som_shepard.pd
- 09.reproducao_audio.pd
- 10.implementacao_audio_equacao_5.pd
- 11.implementacao_audio_equacao_6
- 12.equacao_7_env_pollack.pd
- 13.conversao_env_burns.pd
- 14.conversao_env_nakajima.pd
- 15.equacao_9_env_deutsch.pd
- 16.envolventes_espectrais.pd
- 17.equacao_10_envolvente_espectral_alternativa.pd
- 18.EEAAG.pd
- 19.Chromasom.pd
- 20.frequencias_escalas_24_e_48_partes.pd
- 21.amplitudes_escalas_24_e_48_partes.pd

- 22.series_exercicios.pd
- 23.FM.pd
- 24.chromasom_fm.pd
- 25.chromasom_fm_presets.pd

Directório sons:

- 01.circularity_in_pitch_judgment.wav
- 03.melodia_original.wav
- 04.melodia_preservada.wav
- 05.melodia_aleatoria.wav

Directório vídeos:

- 01.poste_barbeiro.mov
- 02.smc_2012.avi

ANEXO 4: FORMULÁRIO

INTRODUÇÃO

Pedro Luís de Moraes Patrício é Investigador do Centro de Investigação em Ciência e Tecnologia das Artes (CITAR) e Doutorando em Informática Musical pela Escola das Artes (EA), Universidade católica Portuguesa (UCP), Porto, Portugal.

Vem por este meio solicitar a sua colaboração através do preenchimento de um formulário que se encontra representado nas páginas seguintes.

Com o preenchimento do formulário pretendemos avaliar o trabalho experimental que tem vindo a ser desenvolvido no âmbito da nossa tese de doutoramento, cujo título é: **Ilusões Sonoras: erros dos sentidos, verdades da percepção – um estudo sobre a aplicação dos mecanismos das ilusões sonoras na composição musical digital.**

Referimo-nos concretamente à ilusão sonora conhecida pela Ilusão da Escala de Shepard (IES). A IES é uma escala dividida em 12 partes iguais, que quando escutada cria a ilusão sonora de que sobe infinitamente ao longo de uma espiral, como acontece no plano visual com a escada interminável de Escher (1960), (Figura 1).

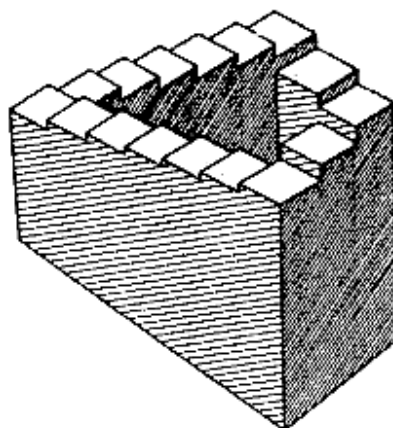


Figura. 1 – Representação simplificada da escada interminável de Escher (1960). Imagem extraída de Shepard (1964, p. 2348).

FORMULÁRIO

Nome: _____ Idade: _____

Área musical: _____ N.º de anos de prática musical _____

Instituição: _____ Data: _____

Após ter escutado a versão original da Ilusão da Escala de Shepard (IES), responda se:

1. o exercício n.º 1 reproduz com sucesso a IES;

SIM ☐

NÃO ☐

2. se no exercício n.º 2, o efeito sonoro da IES se mantém.

SIM ☐

NÃO ☐

Obrigado pela colaboração.

